

# SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB#

86576

Requester's Full Name: Colins Qian Examiner #: 75710 Date: 2/12/03  
 Art Unit: 1636 Phone Number 30 6-0283 Serial Number: 09/758962  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Expression of Foreign genes from plant virus vectors  
 Inventors (please provide full names): Simon Santa-Cruz et al.

Earliest Priority Filing Date: 1/7/2001

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

*Please search SEQID NO: 1, and also do a nucleotide search.*  
*8/19/03*

Edward Hart  
 Technical Info. Specialist  
 STIC/Biotech  
 CMI 6B02 Tel: 305-9203

BEST AVAILABLE COPY

## STAFF USE ONLY

### Type of Search

NA Sequence (#) 1-1 STN Revised to ST Vendors and cost where applicable

AA Sequence (#) \_\_\_\_\_ Dialog \_\_\_\_\_

Structure (#) \_\_\_\_\_ Questel/Orbit \_\_\_\_\_

Bibliographic \_\_\_\_\_ Dr.Link \_\_\_\_\_

Litigation \_\_\_\_\_ Lexis/Nexis \_\_\_\_\_

Fulltext \_\_\_\_\_ Sequence Systems 04

Patent Family \_\_\_\_\_ WWW/Internet \_\_\_\_\_

Other \_\_\_\_\_ Other (specify) \_\_\_\_\_

Searcher: \_\_\_\_\_

Searcher Phone #: \_\_\_\_\_

Searcher Location: \_\_\_\_\_

Date Searcher Picked Up: 2/13/03

Date Completed: 2/19/03

Searcher Prep & Review Time: \_\_\_\_\_

Clerical Prep Time: \_\_\_\_\_

Online Time: \_\_\_\_\_

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2003, 06:08:11 ; Search time 2601 seconds

(without alignments)  
2103.548 Million cell updates/sec

Title: US-09-758-962-1

Perfect score: 188  
Sequence: 1 gtcgacgctacataagct.....gaagaagaagctcaccatg 188

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: GenEmbl:  
1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vtl:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	149	79.3	1537	14 S68164	S68164, transport p
2	149	79.3	6312	14 TMVRP/PCP	Z29370 Tobacco mos
3	42	22.3	6311	14 BRU03387	U03387 Turnip vein
4	38	20.2	1568	14 MDV/PRA	L41561 Gallid herp
5	35	18.6	2126	6 AR067739	AR067739 Sequence
6	35	18.6	2126	6 AR096004	AR096004 Sequence
7	35	18.6	2126	6 AR116595	AR116595 Sequence
8	35	18.6	2126	6 AR121154	AR121154 Sequence
9	35	18.6	7560	6 AR153837	AR153837 Sequence
10	34	18.1	110	10 AB027145	AB027145 Rattus no
11	34	18.1	338	11 CSPAGSR	Z47159 Calothrix D
12	34	18.1	369	11 G16087	Z47159 chicken STS
13	34	18.1	373	11 G01691	G16087 chicken STS
14	34	18.1	397	1 CSPAG13R	Z47173 Calothrix D
15	34	18.1	487	4 HRSTCRAG	L27850 Equus cabal
16	34	18.1	741	6 AX156405	AX156405 Sequence
17	34	18.1	759	6 E14482	AX156405 Sequence
18	34	18.1	924	8 BNPBP3	E14482 cDNA encodi
19	34	18.1	924	8 BNPBP3	A59874 Sequence 3
20	34	18.1	1043	8 CAR011383	X90729 B.napus MRN
21	34	18.1	1097	8 AB049721	AJ011383 Cicer ari
22	34	18.1	1181	6 E14788	AB049721 Pisum sat
23	34	18.1	1221	10 M4TCEBP1	E14788 cDNA encodi
24	34	18.1	1318	6 AX304651	X54499 M.musculus
25	34	18.1	1318	6 E37094	AX304651 Sequence
26	34	18.1	1464	6 AX286689	E37094 GABA BP pol
27	34	18.1	1638	6 AX107064	AB067774 Homo sap1
28	34	18.1	1638	6 AX136732	AX286689 Sequence
29	34	18.1	1638	6 AR169244	AR107064 Sequence
30	34	18.1	1680	5 XLU09632	AR136732 Sequence
31	34	18.1	1845	6 AB085846	AR169244 Sequence
32	34	18.1	1889	6 AX049343	U09632 Xenopus lae
33	34	18.1	1958	6 A94995	AB085846 Chlorella
34	34	18.1	1958	6 AX022639	AX049343 Sequence
35	34	18.1	1958	6 E38366	A94995 Sequence 1
36	34	18.1	1958	6 E39010	AX022639 Sequence
37	34	18.1	1990	8 CCA401353	E38366 Binding par
38	34	18.1	1998	8 CRI401355	E39010 Nucleic ac1
39	34	18.1	2142	8 SOJ00265	AJ401353 Cuscuta p
40	34	18.1	2171	4 AB051103	AJ000265 Spinacia
41	34	18.1	2185	6 AR092424	AB051103 Felis cat
42	34	18.1	2199	3 LME292039	AR092424 Sequence
43	34	18.1	2200	6 AR050076	AJ292039 Leishman1
44	34	18.1	2378	14 AB003936	AR050076 Sequence
45	34	18.1	2378	14 AB003936	AB003936 Crucifer

## ALIGNMENTS

RESULT 1  
S68164  
LOCUS S68164  
DEFINITION transport protein, capsid protein [Tobamovirus, host: cruciferous plants, Genomic, 1537 nt].  
ACCESSION S68164  
VERSION S68164  
KEYWORDS 1537 bp DNA linear VRL 23-SEP-1994  
SOURCE Tobamovirus  
ORGANISM Tobamovirus host: cruciferous plants.  
REFERENCE 1 (bases 1 to 1537)  
Dorokhov, Iu. L., Ivanov, P. A., Novikov, V. K., Yefimov, V. A. and Alabekov, I. G.  
Tobamovirus of cruciferous plants: nucleotide sequence of genes of



YDQKSEFHCVAVYKIMEKIDGIDVETAEYKROGHRKTLKDYTAGICRTQCLMORSSD  
 VTTTIGTITLILACLSMIDMDVIAAECGDDSLIYIKGIDLDIDAGANLMMKPE  
 AKIFRKYGFCEGCVITHHGRGAIYVDPLKLSKIGCHIDVHLELBSLSDVA  
 SNLNCASFQSLDEAVAEVHKTVAGSGFACSLIKLSKRLFRDLEFV"  
 3362                    3363                    3364                    3365                    3366                    3367                    3368                    3369                    3370                    3371                    3372                    3373                    3374                    3375                    3376                    3377                    3378                    3379                    3380                    3381                    3382                    3383                    3384                    3385                    3386                    3387                    3388                    3389                    3390                    3391                    3392                    3393                    3394                    3395                    3396                    3397                    3398                    3399                    3400                    3401                    3402                    3403                    3404                    3405                    3406                    3407                    3408                    3409                    3410                    3411                    3412                    3413                    3414                    3415                    3416                    3417                    3418                    3419                    3420                    3421                    3422                    3423                    3424                    3425                    3426                    3427                    3428                    3429                    3430                    3431                    3432                    3433                    3434                    3435                    3436                    3437                    3438                    3439                    3440                    3441                    3442                    3443                    3444                    3445                    3446                    3447                    3448                    3449                    3450                    3451                    3452                    3453                    3454                    3455                    3456                    3457                    3458                    3459                    3460                    3461                    3462                    3463                    3464                    3465                    3466                    3467                    3468                    3469                    3470                    3471                    3472                    3473                    3474                    3475                    3476                    3477                    3478                    3479                    3480                    3481                    3482                    3483                    3484                    3485                    3486                    3487                    3488                    3489                    3490                    3491                    3492                    3493                    3494                    3495                    3496                    3497                    3498                    3499                    3500                    3501                    3502                    3503                    3504                    3505                    3506                    3507                    3508                    3509                    3510                    3511                    3512                    3513                    3514                    3515                    3516                    3517                    3518                    3519                    3520                    3521                    3522                    3523                    3524                    3525                    3526                    3527                    3528                    3529                    3530                    3531                    3532                    3533                    3534                    3535                    3536                    3537                    3538                    3539                    3540                    3541                    3542                    3543                    3544                    3545                    3546                    3547                    3548                    3549                    3550                    3551                    3552                    3553                    3554                    3555                    3556                    3557                    3558                    3559                    3560                    3561                    3562                    3563                    3564                    3565                    3566                    3567                    3568                    3569                    3570                    3571                    3572                    3573                    3574                    3575                    3576                    3577                    3578                    3579                    3580                    3581                    3582                    3583                    3584                    3585                    3586                    3587                    3588                    3589                    3590                    3591                    3592                    3593                    3594                    3595                    3596                    3597                    3598                    3599                    3600                    3601                    3602                    3603                    3604                    3605                    3606                    3607                    3608                    3609                    3610                    3611                    3612                    3613                    3614                    3615                    3616                    3617                    3618                    3619                    3620                    3621                    3622                    3623                    3624                    3625                    3626                    3627                    3628                    3629                    3630                    3631                    3632                    3633                    3634                    3635                    3636                    3637                    3638                    3639                    3640                    3641                    3642                    3643                    3644                    3645                    3646                    3647                    3648                    3649                    3650                    3651                    3652                    3653                    3654                    3655                    3656                    3657                    3658                    3659                    3660                    3661                    3662                    3663                    3664                    3665                    3666                    3667                    3668                    3669                    3670                    3671                    3672                    3673                    3674                    3675                    3676                    3677                    3678                    3679                    3680                    3681                    3682                    3683                    3684                    3685                    3686                    3687                    3688                    3689                    3690                    3691                    3692                    3693                    3694                    3695                    3696                    3697                    3698                    3699                    3700                    3701                    3702                    3703                    3704                    3705                    3706                    3707                    3708                    3709                    3710                    3711                    3712                    3713                    3714                    3715                    3716                    3717                    3718                    3719                    3720                    3721                    3722                    3723                    3724                    3725                    3726                    3727                    3728                    3729                    3730                    3731                    3732                    3733                    3734                    3735                    3736                    3737                    3738                    3739                    3740                    3741                    3742                    3743                    3744                    3745                    3746                    3747                    3748                    3749                    3750                    3751                    3752                    3753                    3754                    3755                    3756                    3757                    3758                    3759                    3760                    3761                    3762                    3763                    3764                    3765                    3766                    3767                    3768                    3769                    3770                    3771                    3772                    3773                    3774                    3775                    3776                    3777                    3778                    3779                    3780                    3781                    3782                    3783                    3784                    3785                    3786                    3787                    3788                    3789                    3790                    3791                    3792                    3793                    3794                    3795                    3796                    3797                    3798                    3799                    3800          

[illegible]

```

codon_start=1
product="29k"
protein_id="CA62560.1"
db_xref="SPTREMBL:Q88605"
db_xref="SPTREMBL:gi:488715"
/translation="MSIYSYEPKVDSEPLNLSKMEILPKALTRKTVYSISTDIISVK
ESSTLEIDIDLINPLDMDRYRYGLGAVFTKRPQWYHRYRDLKIEGWPLLEYVS
VICTYRAAKSKRFQKLPNVEYSTVPAKRPQWYHRYRDLKIEGWPLLEYVS
VAVTNNVWVKGRKREKVAINDPVEGEGVADDEVDSVAAFKADNFKRKRVK
GVYSKIYKREKAGDSENLKEENYLQIKESVPERSGVGAHSDA"
5604..6077
/function="coat protein"

```

1800 a

[illegible]

	Matches	143,	5514
CGAATTCGTCGATTGGCTGCAGCATTAAAGCGGTGACACCTTAAAGCAAAA			

[illegible]

87 GAAGCTTGAAGGATTTTCCTAAGTATACTACAGACC GGAGAAGTTC

Db : 5515 GAAGGAGGAA 175

QY

Db 53/3 1000000

Result	Size (bp)	Genome
RESULT 3	6311 bp	RNA
BRU03387		complete genome.

DEFINITION	
U03387	L22518
U03387.1	GI:2894629

Turnip vein-clearing virus, no DNA stage  
Turnip vein-clearing viruses, no DNA stage  
Turnip vein-clearing viruses; ssRNA positive-strand viruses; ssRNA positive-strand viruses

REFERENCE	AUTHORS	TITLE
2	Larley, R. T., Lane, D. C.	Electron microscopic and molecular cloning of the genome of the hepatitis A virus

JOURNAL ARCH. VICE.  
95091521  
MEDLINE  
7008835  
3003 TO 6311)

REFERENCE  
AUTHORS  
TITLE  
2 Voss, J. C.  
Larley, R. T.,  
Completion of a cDNA seq

JOURNAL  
Gene 100, 221-226 (1982)  
3 (bases 1 to 6311)  
Mol. Biol. Evol. 1:1-10, 1984.  
© 1984 by the University of Chicago  
0895-9135/84/010001-10\$02.00/0  
DOI: 10.1093/molbev/1.1.1  
Printed in the USA

Submitted (09-NOV-2005)  
NRC 246, Stillwater, MN  
State University, MN 55057-1063

Melchner, V. *Biochemistry & Molecular Biology*, 25, 125-130 (1998)  
 Direct Submission  
 Submitted (18-FEB-1998) Blochemistry & Molecular Biology, OK 74078-0454, USA  
 and 746. Stillwater,  
 TITLE

	REMARK	Sequence update by version ref-
	REMARK	On Feb 18, 1998 this sequence
	COMMENT	recognition/Qualifiers

Source

**CDS**

```

tyrosine start=1
...3389...3391,aa:Tyr)
...component'

```

/protein\_id="AAC02130"  
/protein="GT:619645"

RPKVNHS...  
STYDIGNESAHLEKGRD

IKYVCKTEPASÜKE V ...

AMIEE...  
KVAEOALEIKVPELYCTE...  
...OUN

KGTMSAVYTGSLNVQ...

КМЛІККАН:О  
ОСНВАУГДТКОІРЕІО

[illegible]



```

FEATURES
  SOURCE      Location/Qualifiers
  BASE COUNT  317 a 763 c 669 g 377 t
  ORIGIN
Query Match      18.6%; Score 35; DB 6; Length 2126;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 35
Db 30 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 64

RESULT 7
LOCUS      AR116595      2126 bp      DNA      linear      PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6133420.
ACCESSION  AR116595
VERSION     AR116595.1 GI:14096917
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 2126)
AUTHORS     Ames,R.S., Jr., Sarau,H.M., Foley,J.J., Shabon,U., Bergsma,D. and
            Chambers,J.K.
TITLE       GPR14 polypeptides
JOURNAL     Patent: US 6133420-A 1 17-OCT-2000;
FEATURES
  SOURCE      Location/Qualifiers
  BASE COUNT  317 a 763 c 669 g 377 t
  ORIGIN
Query Match      18.6%; Score 35; DB 6; Length 2126;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 35
Db 30 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 64

RESULT 8
LOCUS      AR121154      2126 bp      DNA      linear      PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6159700.
ACCESSION  AR121154
VERSION     AR121154.1 GI:14104730
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 2126)
AUTHORS     Alyar,N.Y., Ames,R.S., Arnold,A., Romanic., Al-Barazani,I.K.,
            Bergsma,D.J., Chambers,J., Douglas,S.A., Foley,J.J., Gout,B.,
            Khandoudi,N., Sarau,H.M., Shabon,U. and Willette,R.N.
TITLE       Method of finding agonist and antagonist to human and rat GPR14
JOURNAL     Patent: US 6159700-A 1 12-DEC-2000;
FEATURES
  SOURCE      Location/Qualifiers
  BASE COUNT  317 a 763 c 669 g 377 t
  ORIGIN
Query Match      18.6%; Score 35; DB 6; Length 2126;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 35

```

```

Db 30 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 64

RESULT 9
LOCUS      AR153837      7560 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION Sequence 4 from patent US 6235975.
ACCESSION  AR153837
VERSION     AR153837.1 GI:15121369
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 7560)
AUTHORS     Harada,J.J., Lotan,T., Ohno,M.-a., Goldberg,R.B. and Fischer,R.L.
TITLE       Leafy cotyledon1 genes and methods of modulating embryo development
JOURNAL     Patent: US 6235975-A 4 22-MAY-2001;
FEATURES
  SOURCE      Location/Qualifiers
  BASE COUNT  2378 a 1326 c 1264 g 2573 t 19 others
  ORIGIN
Query Match      18.6%; Score 35; DB 6; Length 7560;
Best Local Similarity 100.0%; Pred. No. 9.8e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 35
Db 53 GTCGACGGTATCGATAGCTTGATATCGAATTGCT 87

RESULT 10
LOCUS      AB027145      110 bp      mRNA      linear      ROD 06-JUL-2000
DEFINITION Rattus norvegicus mRNA for CDcrel-1f, partial cds.
ACCESSION  AB027145
VERSION     AB027145.1 GI:8953678
KEYWORDS   CDcrel-1/PNUTL1 fetus form.
SOURCE      Rattus norvegicus (strain:Wistar) neonatal male neocortex cDNA to
            mRNA.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine;
            Rattus.
REFERENCE    1 (sites)
AUTHORS     Toda,S., Kajii,Y., Sato,M. and Nishikawa,T.
TITLE       Reciprocal expression of infant- and adult-prefering transcripts
            of CDcrel-1 septin gene in the rat neocortex
JOURNAL     Biochem. Biophys. Res. Commun. 273 (2), 723-728 (2000)
MEDLINE     20334336
REFERENCE    2 (bases 1 to 110)
AUTHORS     Toda,S., Kajii,Y. and Nishikawa,T.
TITLE       Direct Submission
JOURNAL     Submitted (11-MAY-1999) Shigenobu Toda, National Institute of
            Neuroscience, NCNP, Department of Mental Disorder Research; 4-1-1
            Ogawabagashi, Kodaira, Tokyo 187-8502, Japan
            (E-mail:toda@psy.med.tohoku.ac.jp, Tel:81-423-46-1714,
            Fax:81-423-46-1744)
FEATURES
  SOURCE      Location/Qualifiers
  BASE COUNT  110
  ORIGIN
  /organism="Rattus norvegicus"
  /strain="Wistar"
  /db_xref="taxon:10116"
  /sex="male"
  /tissue_type="neocortex"
  /dev_stage="neonatal"
  1..110
  /gene="CDcrel-1/PNUTL1f"
  57..>110
  /gene="CDcrel-1/PNUTL1f"
  /codon_start=1

```

/product="CDCrel-1F"  
 /protein\_id="BA98052.1"  
 /db\_xref="GI:8953679"  
 /translation="MSTGLRKSLAPEDKQ"

BASE COUNT 30 a 31 c 31 g 18 t  
 ORIGIN

Query Match 18.1%; Score 34; DB 10; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGCAGGTATCGATAGCTTGATATCGAATTGC 34  
 |||||||  
 Db 20 GTGCAGGTATCGATAGCTTGATATCGAATTGC 53

RESULT 11  
 CSPA5R 338 bp DNA Linear BCT 04-JAN-1995  
 LOCUS Calothrix D253 genomic DNA (clone AG5R).  
 DEFINITION 247159 GI:619119  
 ACCESSION 247159.1 GI:619119  
 VERSION  
 KEYWORDS  
 SOURCE Calothrix sp.  
 ORGANISM Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Calothrix.

REFERENCE  
 AUTHORS Robinson, N.J., Robinson, P.J., and Gupta, A.  
 TITLE Genomic DNA sequence obtained in the process of generating a random  
 sequence database greater than 10 kb from single sequence runs from  
 random clones of a Calothrix D253 library  
 JOURNAL Unpublished

REFERENCE  
 AUTHORS Robinson, N.J., Robinson, P.J., Gupta, A., Bleasby, A.J., Whitton, B.A.  
 and Morby, A.P.  
 TITLE Singular over-representation of HPI1 in DNA from many cyanobacteria  
 JOURNAL Unpublished

REFERENCE  
 AUTHORS Robinson, N.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-DEC-1994) Nigel J. Robinson, Biochemistry and  
 Genetics, The Medical School, University of Newcastle, Framlington  
 Place, Newcastle upon Tyne, NE2 4HH, UK  
 4 (bases 1 to 338)  
 Robinson, N.J., Robinson, P.J., Gupta, A., Bleasby, A.J., Whitton, B.A.  
 and Morby, A.P.  
 TITLE Singular over-representation of an octameric palindrome, HPI1, in  
 DNA from many cyanobacteria  
 JOURNAL Nucleic Acids Res. 23 (5), 729-735 (1995)  
 PUBMED 95223777  
 FEATURES 7708486

FEATURES  
 source  
 1..338  
 Location/Qualifiers  
 /organism="Calothrix sp."  
 /strain="D253"  
 /db\_xref="taxon:1187"  
 /clone="AG5R"  
 /clone\_11b="LZAPCalothrix-LIB"  
 BASE COUNT 83 a 87 c 77 g 91 t  
 ORIGIN

Query Match 18.1%; Score 34; DB 1; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGCAGGTATCGATAGCTTGATATCGAATTGC 34  
 |||||||  
 Db 138 GTGCAGGTATCGATAGCTTGATATCGAATTGC 105

RESULT 12  
 G16087/c 369 bp DNA Linear STS 13-FEB-1996  
 LOCUS

DEFINITION chicken STS ADL220, sequence tagged site.  
 ACCESSION G16087  
 VERSION G16087.1 GI:1184628  
 STS: STS sequence; primer: sequence tagged site.  
 KEYWORDS Gallus gallus vector-plu1script II KS+ host-E. coli.  
 SOURCE Gallus gallus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianidae; Gallus.  
 1 (bases 1 to 369)  
 Cheng, H.H.  
 REFERENCE Unpublished (1994)  
 AUTHORS  
 JOURNAL  
 COMMENT

USDA-ARS  
 3606 E. Mount Hope Rd, East Lansing, MI 48823, USA  
 Tel: 5173376758  
 Fax: 5173376776  
 Email: hcheng@pilot.msu.edu

Primer A: AGGGGTATTGCTGCTGC  
 Primer B: GTCCTCAGTCGCAATGC  
 STS size: 114  
 PCR Profile:

Denaturation: 94 degrees C for 1 minute  
 Annealing: 51 degrees C for 1 minute  
 Polymerization: 72 degrees C for 1 minute  
 PCR Cycles: 30  
 Thermal Cycler: MJ Research

Protocol:  
 Template: 50-100 ng  
 Primer: 0.1 uM  
 dNTPs: 200 uM  
 Taq polymerase: 0.4 units/ul  
 Total Vol: 25 ul

Buffer:  
 MgCl2: 1.5 mM  
 KCl: 50 mM  
 Tris-HCl: 10 mM  
 pH: 9.0  
 Triton X-100: 0.1 %  
 Location/Qualifiers  
 1..369  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"

FEATURES  
 source  
 STS  
 primer\_bind 35..148  
 primer\_bind 35..54  
 BASE COUNT 96 a 54 c 112 g 107 t  
 ORIGIN

Query Match 18.1%; Score 34; DB 11; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGCAGGTATCGATAGCTTGATATCGAATTGC 34  
 |||||||  
 Db 323 GTGCAGGTATCGATAGCTTGATATCGAATTGC 290

RESULT 13  
 G01691/c 373 bp DNA Linear STS 07-DEC-1994  
 LOCUS  
 DEFINITION chicken STS ADL271, sequence tagged site.  
 ACCESSION G01691  
 VERSION G01691.1 GI:595215  
 STS: STS sequence; primer: sequence tagged site.  
 KEYWORDS Gallus gallus vector-plu1script II KS+ host-E. coli.  
 SOURCE Gallus gallus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 373)  
 Phasianinae: Gallus.  
 AUTHORS Cheng, H.H.  
 JOURNAL Unpublished (1994)  
 COMMENT Synonyms: B367  
 Contact: Hans H. Cheng  
 Avian Disease and Oncology Laboratory  
 USDA-ARS  
 3606 E. Mount Hope Rd, East Lansing, MI 48823, USA  
 Tel: 5173376758  
 Fax: 5173376776  
 Email: hcheng@pilot.msu.edu

Primer A: TCATCAGAACCCCAAGCACA  
 Primer B: ATGAATGAACCATCTAAC  
 STS size: 136

PCR Profile:  
 Denaturation: 94 degrees C for 1 minute  
 Annealing: 48 degrees C for 1 minute  
 Polymerization: 72 degrees C for 1 minute  
 PCR Cycles: 30  
 Thermal Cycler: MJ Research

Protocol:  
 Template: 50-100 ng  
 Primer: 0.1 uM  
 dNTPs: 200 uM  
 Taq polymerase: 0.4 units/ul  
 Total Vol: 25 ul

Buffer:  
 MgCl2: 1.5 mM  
 KCl: 50 mM  
 Tris-HCl: 10 mM  
 pH: 9.0  
 Triton X-100: 0.1 %

Location/Qualifiers  
 1..373  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"

STS  
 primer\_bind 3..22  
 primer\_bind complement(119..138) 113 t

BASE COUNT 98 a 81 c 81 g 113 t

ORIGIN

Query Match 18.1%; Score 34; DB 11; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGCTATCGATAGCTTGATATCGAATTCG 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 257 GTCGACGCTATCGATAGCTTGATATCGAATTCG 224

RESULT 14  
 CSPAG13R/C 397 bp DNA linear BCT 04-JAN-1995  
 DEFINITION Calothrix D253 genomic DNA (clone AG13R).  
 ACCESSION 247173  
 VERSION 247173.1 GI:619086  
 KEYWORDS  
 SOURCE Calothrix sp.  
 ORGANISM Calothrix sp.  
 Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Calothrix.  
 1 (bases 1 to 397)  
 Robinson, N.J., Robinson, P.J. and Gupta, A.  
 Genomic DNA sequence obtained in the process of generating a random  
 sequence database of a Calothrix D253 library  
 Random clones of a Calothrix D253 library  
 Unpublished  
 2 (bases 1 to 397)  
 Robinson, N.J., Robinson, P.J., Gupta, A., Bleasby, A.J., Whitton, B.A.  
 and Morby, A.F.

JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE

TITLE Singular over-representation of H1P1 in DNA from many cyanobacteria  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 397)  
 AUTHORS Robinson, N.J.  
 JOURNAL Direct Submission  
 Submitted (16-DEC-1994) Nigel J. Robinson, Biochemistry and  
 Genetics, The Medical School, University of Newcastle, Framlington  
 Place, Newcastle upon Tyne, NE2 4HH, UK  
 4 (bases 1 to 397)  
 Robinson, N.J., Robinson, P.J., Gupta, A., Bleasby, A.J., Whitton, B.A.  
 and Morby, A.F.  
 Singular over-representation of an octameric palindrome, H1P1, in  
 DNA from many cyanobacteria  
 Nucleic Acids Res. 23 (5), 729-735 (1995)  
 9523777  
 7708486

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED

FEATURES  
 source  
 Location/Qualifiers  
 1..397  
 /organism="Calothrix sp."  
 /strain="D253"  
 /db\_xref="taxon:1187"  
 /clone="AG13R"  
 /clone\_11b="L2APCalothrix-L1b"

BASE COUNT 105 a 91 c 88 g 113 t

ORIGIN

Query Match 18.1%; Score 34; DB 1; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGCTATCGATAGCTTGATATCGAATTCG 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 90 GTCGACGCTATCGATAGCTTGATATCGAATTCG 57

RESULT 15  
 HRS1CRAG 487 bp DNA linear MAM 01-MAY-1995  
 LOCUS Equus caballus (clone T131) T-cell receptor DNA, V-region.  
 DEFINITION L27850  
 ACCESSION L27850.1 GI:790760  
 VERSION L27850.1  
 KEYWORDS T-cell receptor; V-region.  
 SOURCE Equus caballus (clone: T131) DNA.  
 ORGANISM Equus caballus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 1 (bases 1 to 487)  
 Schrenzel, M.D., Watson, J.L. and Ferrick, D.A.  
 Characterization of horse (Equus caballus) T-cell receptor beta  
 chain genes  
 Immunogenetics 40 (2), 135-144 (1994)  
 94299284  
 7913080

JOURNAL  
 MEDLINE  
 PUBMED

FEATURES  
 source  
 Location/Qualifiers  
 1..487  
 /organism="Equus caballus"  
 /db\_xref="taxon:9796"  
 /clone="T131"  
 <1..>430  
 /note="This CDS feature is included to show the  
 translation of the corresponding V-region. Presently  
 translation qualifiers on V-region features are illegal."  
 /codon\_start=2  
 /protein\_id="AA65661.1"  
 /db\_xref="GI:790761"  
 /translation="SRSTVISLSISNSARGHISLSEGLVSKPLMLFLLLGPTGS  
 GLGVLSOHSRAICKNSVEMECRTYDLQVCMYQVNAQNGITPVANSNGLLSS  
 TYEQQFVKKFPIPSHNLFTSLITLCAHPNSGLIFSCAR"  
 <1..430  
 /product="T cell receptor"  
 /note="DJ"

V\_region  
 J\_segment  
 BASE COUNT 114 a 132 c 123 g 118 t



## ORIGIN

Query Match 18.1%; Score 34; DB 4; Length 487;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGCTATCGATTAAGCTGTGATATCGAATTCG 34  
|||||  
DB 7 GTCGACGCTATCGATTAAGCTGTGATATCGAATTCG 40

Search completed: February 19, 2003, 07:57:01  
Job time : 2608 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2003, 06:07:16 ; Search time 229 Seconds  
(without alignments)  
1848.802 Million cell updates/sec

Title: US-09-758-962-1  
Perfect score: 188  
Sequence: 1 gtcacgcatcgcataagct.....gaagaagcctcaccatcg 188

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	100.0	188	24	ABQ76058
2	181	96.3	181	24	ABQ76062
3	149	79.3	187	24	ABQ76060
4	149	79.3	251	24	ABQ76059
5	44	23.4	180	24	ABQ76063
6	43	22.9	181	24	ABQ76064
7	35	18.6	122	24	ABQ76061
8	35	18.6	444	23	ABY36222
9	35	18.6	444	23	ABY45268
					Human prostate exp
					Human prostate exp

C	10	35	18.6	514	23	ABV49873	Human prostate exp
C	11	35	18.6	1353	24	ABK52453	Accessory factor T
C	12	35	18.6	2126	19	AAV41999	Human GPR14 (G-pro
C	13	35	18.6	2126	20	AAZ11174	Human GPR14 coding
C	14	35	18.6	3236	24	AAK99298	CDNA of the human
C	15	35	18.6	3905	24	ABN83030	Human X2 theca rep
C	16	35	18.6	7560	19	AAV61024	Arabidopsis leafy-7.4 kb Exptl fragm
C	17	35	18.6	4424	22	AAZ45840	Leafy cotyledon 1
C	18	35	18.6	7560	22	AAO16916	Plasmid NBE-T frag
C	19	34	18.1	99	23	ABL53457	Human nervous syst
C	20	34	18.1	353	22	ABAI1976	Human nervous syst
C	21	34	18.1	361	21	AAK59649	Human secreted pro
C	22	34	18.1	362	21	AAFI6468	Human secreted can
C	23	34	18.1	420	20	AAK37517	Human secreted pro
C	24	34	18.1	421	22	ABAI3348	Human nervous syst
C	25	34	18.1	470	20	AAK37520	Human secreted pro
C	26	34	18.1	491	22	ABAI3347	Human nervous syst
C	27	34	18.1	501	21	ABAI4279	Human nervous syst
C	28	34	18.1	502	21	AAK59628	Human secreted pro
C	29	34	18.1	508	22	ABAI3368	Human nervous syst
C	30	34	18.1	523	22	ABAI3365	Human nervous syst
C	31	34	18.1	528	22	ABAI4286	Human nervous syst
C	32	34	18.1	585	22	ABAI3335	Human nervous syst
C	33	34	18.1	587	20	AAK37393	Human secreted pro
C	34	34	18.1	591	24	ABN90021	Mouse clone IMX3-6
C	35	34	18.1	602	22	ABAI3360	Human nervous syst
C	36	34	18.1	607	22	ABAI4282	Human nervous syst
C	37	34	18.1	610	22	ABAI3372	Human nervous syst
C	38	34	18.1	616	20	AAK2262	Human secreted pro
C	39	34	18.1	626	22	AAK41400	CDNA encoding nove
C	40	34	18.1	642	22	ABAI3400	Human nervous syst
C	41	34	18.1	655	22	ABAI4170	Human nervous syst
C	42	34	18.1	712	22	AAO05161	Human secreted pro
C	43	34	18.1	723	22	ABAI4285	Human nervous syst
C	44	34	18.1	740	22	AAK43067	Human secreted pro
C	45	34	18.1	741	22	AAH5734	Human breast tumor

#### ALIGNMENTS

RESULT 1  
ABQ76058 standard; DNA; 188 BP.

AC ABQ76058;  
DT 30-SEP-2002 (first entry)  
DE Clone TXS.GFP-IRES-CP containing IRES DNA sequence.  
XX  
XX  
XX IRES: Internal ribosome entry site; viral protein; bicistronic;  
KW virus infection; coat protein; protein replacement therapy; crop;  
KM nutritional value; seed oil content; ds.  
XX  
XX  
XX Synthetic.  
OS  
PN MO200255719-A2.  
PD 18-JUL-2002.  
XX  
XX 09-JAN-2002; 2002MO-US01123.  
PF  
XX 09-JAN-2001; 2001US-0758962.  
PR  
XX (BIOS-) BIOSOURCE GENETICS CORP.  
PA Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;  
XX WPI; 2002-557829/59.  
XX  
XX  
XX New polypeptides contained in plant virus expression vectors, as gene  
PT expression tools, in protein replacement therapy or for intervening in

PT a metabolic pathway to improve the nutritional value of a crop or alter  
 PT the oil content of seeds  
 XX  
 PS Claim 4; Fig 3; 33pp; English.  
 XX  
 CC This invention describes a novel isolated polynucleotide comprising an  
 CC internal ribosome entry site (IRES) nucleotide sequence, an open reading  
 CC frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,  
 CC where the IRES is located between ORF1 and ORF2. The novel  
 CC polynucleotides of the invention are used in the construction of a  
 CC recombinant potato virus X-based viral vector containing a nucleic acid  
 CC construct comprising a bicistronic message with an intervening IRES. The  
 CC constructs are used in a method for regulating the rate at which a virus  
 CC infection spreads in a host. Regulation is achieved by placing the  
 CC nucleic acid construct comprising an internal ribosome entry site  
 CC upstream of a coat protein gene, where the IRES is chosen by the rate of  
 CC infection of the viral vector on a host in the presence of that IRES. The  
 CC polynucleotide and vectors of the invention are useful for directing  
 CC rapid and high-level expression of foreign genes in mature,  
 CC differentiated, plant tissue. These are particularly useful in protein  
 CC replacement therapy, or for intervening in a metabolic pathway to improve  
 CC the nutritional value of a crop or alter the oil content of the seed.  
 CC This sequence represents a fragment of the clone TKS.GFP.IRESs-CP  
 CC containing an IRES region described in the method of the invention.  
 XX  
 SO Sequence 188 BP; 66 A; 26 C; 51 G; 45 T; 0 other;  
 Query Match 100.0%; Score 188; DB 24; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-81;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTGACGGTATCGATTAAGCTTGATATCGAATTCGTCGATTCGGTTCACGATTAAAGCG 60  
 DB 1 GTGACGGTATCGATTAAGCTTGATATCGAATTCGTCGATTCGGTTCACGATTAAAGCG 60  
 QY 61 GTTGACACTTTAAAG 120  
 DB 61 GTTGACACTTTAAAG 120  
 QY 121 AAGTACAGACCGAGCT 180  
 DB 121 AAGTACAGACCGAGCT 180  
 QY 181 CACCATGG 188  
 DB 181 CACCATGG 188  
 RESULT 2  
 ID ABQ76062 standard; DNA; 181 BP.  
 XX  
 AC ABQ76062;  
 XX  
 DT 30-SEP-2002 (first entry)  
 XX  
 DE Clone TKS.GFP-IRESs-CP containing IREScp DNA sequence.  
 XX  
 OS IRES; internal ribosome entry site; viral protein; bicistronic;  
 KW virus infection; coat protein; protein replacement therapy; crop;  
 XX nutritional value; seed oil content; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200255719-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 09-JAN-2002; 2002WO-US01123.  
 XX  
 PR 09-JAN-2001; 2001US-0758962.  
 XX  
 PA (BIOS-) BIOSOURCE GENETICS CORP.

XX  
 PI Santa Cruz S, Pogue GP, Toth RL, Chapman S, Carr E;  
 XX  
 DR WPI; 2002-557829/59.  
 XX  
 PT New polypeptides contained in plant virus expression vectors, as gene  
 PT expression tools, in protein replacement therapy or for intervening in  
 PT a metabolic pathway to improve the nutritional value of a crop or alter  
 PT the oil content of seeds  
 XX  
 PS Claim 4; Fig 4; 33pp; English.  
 XX  
 CC This invention describes a novel isolated polynucleotide comprising an  
 CC internal ribosome entry site (IRES) nucleotide sequence, an open reading  
 CC frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,  
 CC where the IRES is located between ORF1 and ORF2. The novel  
 CC polynucleotides of the invention are used in the construction of a  
 CC recombinant potato virus X-based viral vector containing a nucleic acid  
 CC construct comprising a bicistronic message with an intervening IRES. The  
 CC constructs are used in a method for regulating the rate at which a virus  
 CC infection spreads in a host. Regulation is achieved by placing the  
 CC nucleic acid construct comprising an internal ribosome entry site  
 CC upstream of a coat protein gene, where the IRES is chosen by the rate of  
 CC infection of the viral vector on a host in the presence of that IRES. The  
 CC polynucleotide and vectors of the invention are useful for directing  
 CC rapid and high-level expression of foreign genes in mature,  
 CC differentiated, plant tissue. These are particularly useful in protein  
 CC replacement therapy, or for intervening in a metabolic pathway to improve  
 CC the nutritional value of a crop or alter the oil content of the seed.  
 CC This sequence represents a fragment of the clone TKS.GFP.IRESs-CP  
 CC containing an IREScp region described in the method of the invention.  
 XX  
 SO Sequence 181 BP; 64 A; 24 C; 49 G; 44 T; 0 other;  
 Query Match 96.3%; Score 181; DB 24; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-78;  
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTGACGGTATCGATTAAGCTTGATATCGAATTCGTCGATTCGGTTCACGATTAAAGCG 60  
 DB 1 GTGACGGTATCGATTAAGCTTGATATCGAATTCGTCGATTCGGTTCACGATTAAAGCG 60  
 QY 61 GTTGACACTTTAAAG 120  
 DB 61 GTTGACACTTTAAAG 120  
 QY 121 AAGTACAGACCGAGCT 180  
 DB 121 AAGTACAGACCGAGCT 180  
 QY 181 C 181  
 DB 181 C 181  
 RESULT 3  
 ID ABQ76060/c  
 XX  
 AC ABQ76060 standard; DNA; 187 BP.  
 XX  
 DT 30-SEP-2002 (first entry)  
 XX  
 DE Clone TKS.GFP-SERI-CP containing IRES DNA sequence.  
 XX  
 OS IRES; internal ribosome entry site; viral protein; bicistronic;  
 KW virus infection; coat protein; protein replacement therapy; crop;  
 XX nutritional value; seed oil content; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200255719-A2.

PD 18-JUL-2002.  
XX  
PF 09-JAN-2002; 2002MO-US01123.  
XX  
PR 09-JAN-2001; 2001US-0758962.  
XX  
PA (BIOS-) BIOSOURCE GENETICS CORP.  
XX  
PI Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;  
XX WPI; 2002-557829/59.  
XX  
PT New polypeptides contained in plant virus expression vectors, as gene  
PT expression tools, in protein replacement therapy or for intervening in  
PT a metabolic pathway to improve the nutritional value of a crop or alter  
PT the oil content of seeds  
XX  
PS Claim 4; Fig 3; 33pp; English.  
XX  
CC This invention describes a novel isolated polynucleotide comprising an  
CC internal ribosome entry site (IRES) nucleotide sequence, an open reading  
CC frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,  
CC where the IRES is located between ORF1 and ORF2. The novel  
CC polynucleotides of the invention are used in the construction of a  
CC recombinant potato virus X-based viral vector containing a nucleic acid  
CC constructs are used in a method for regulating the rate at which a virus  
CC infection spreads in a host. Regulation is achieved by placing the  
CC nucleic acid construct comprising an internal ribosomes entry site  
CC upstream of a coat protein gene, where the IRES is chosen by the rate of  
CC infection of the viral vector on a host in the presence of that IRES. The  
CC polynucleotide and vectors of the invention are useful for directing  
CC rapid and high-level expression of foreign genes in mature,  
CC differentiated, plant tissue. These are particularly useful in protein  
CC replacement therapy, or for intervening in a metabolic pathway to improve  
CC the nutritional value of a crop or alter the oil content of the seed.  
CC This sequence represents a fragment of the clone TXS.GFP.SERI-CP  
CC containing an IRES region described in the method of the invention.  
XX  
SQ Sequence 187 BP; 45 A; 47 C; 27 G; 68 T; 0 other;  
XX  
Query Match 79.3%; Score 149; DB 24; Length 187;  
Best Local Similarity 100.0%; Pred. No. 1.1e-62;  
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 29 AATTCGCGATTCGGTTCAGCATTTAAAGCGGTGACACTTTAAAGAGAAAGA 88  
DB 180 AATTCGCGATTCGGTTCAGCATTTAAAGCGGTGACACTTTAAAGAGAAAGA 121  
QY 89 AGGTTGAAGAAAGGAGTGTAGTAAAGTATAGTACAGACCGAGAGATACCGCGTC 148  
DB 120 AGGTTGAAGAAAGGAGTGTAGTAAAGTATAGTACAGACCGAGAGATACCGCGTC 61  
QY 149 CTGATTCGTTTAAATTGAAAGAAAGA 177  
DB 60 CTGATTCGTTTAAATTGAAAGAAAGA 32  
XX  
RESULT 4  
AB076059  
ID AB076059 standard; DNA; 251 BP.  
XX  
AC AB076059;  
XX  
DT 30-SEP-2002 (first entry)  
XX  
DE Clone TXS.GFP-HIRES-CP containing IRES DNA sequence.  
XX  
KW IRES: internal ribosome entry site; viral protein; bicistronic;  
KW virus infection; coat protein; protein replacement therapy; crop;  
XX nutritional value; seed oil content; ds.  
XX  
OS Synthetic.

XX  
PN W0200255719-A2.  
XX  
PD 18-JUL-2002.  
XX  
PF 09-JAN-2002; 2002MO-US01123.  
XX  
PR 09-JAN-2001; 2001US-0758962.  
XX  
PA (BIOS-) BIOSOURCE GENETICS CORP.  
XX  
PI Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;  
XX WPI; 2002-557829/59.  
XX  
PT New polypeptides contained in plant virus expression vectors, as gene  
PT expression tools, in protein replacement therapy or for intervening in  
PT a metabolic pathway to improve the nutritional value of a crop or alter  
PT the oil content of seeds  
XX  
PS Claim 4; Fig 3; 33pp; English.  
XX  
CC This invention describes a novel isolated polynucleotide comprising an  
CC internal ribosome entry site (IRES) nucleotide sequence, an open reading  
CC frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,  
CC where the IRES is located between ORF1 and ORF2. The novel  
CC polynucleotides of the invention are used in the construction of a  
CC recombinant potato virus X-based viral vector containing a nucleic acid  
CC constructs are used in a method for regulating the rate at which a virus  
CC infection spreads in a host. Regulation is achieved by placing the  
CC nucleic acid construct comprising an internal ribosomes entry site  
CC upstream of a coat protein gene, where the IRES is chosen by the rate of  
CC infection of the viral vector on a host in the presence of that IRES. The  
CC polynucleotide and vectors of the invention are useful for directing  
CC rapid and high-level expression of foreign genes in mature,  
CC differentiated, plant tissue. These are particularly useful in protein  
CC replacement therapy, or for intervening in a metabolic pathway to improve  
CC the nutritional value of a crop or alter the oil content of the seed.  
CC This sequence represents a fragment of the clone TXS.GFP.HIRES-CP  
CC containing an IRES region described in the method of the invention.  
XX  
SQ Sequence 251 BP; 72 A; 51 C; 76 G; 52 T; 0 other;  
XX  
Query Match 79.3%; Score 149; DB 24; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-62;  
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 27 CGAATTCGTCGATTCGGTTCAGCATTTAAAGCGGTGACACTTTAAAGAGAAAGA 86  
DB 97 CGAATTCGTCGATTCGGTTCAGCATTTAAAGCGGTGACACTTTAAAGAGAAAGA 156  
QY 87 GAAGGTTGAAGAAAGGAGTGTAGTAAAGTATAGTACAGACCGAGAGATACCGCGG 146  
DB 157 GAAGGTTGAAGAAAGGAGTGTAGTAAAGTATAGTACAGACCGAGAGATACCGCGG 216  
QY 147 TCCTGATTCGTTTAAATTGAAAGAAAGA 175  
DB 217 TCCTGATTCGTTTAAATTGAAAGAAAGA 245  
XX  
RESULT 5  
AB076063  
ID AB076063 standard; DNA; 180 BP.  
XX  
AC AB076063;  
XX  
DT 30-SEP-2002 (first entry)  
XX  
DE Clone SC196 containing IREScp DNA sequence.  
XX  
KW IRES: internal ribosome entry site; viral protein; bicistronic;  
KW virus infection; coat protein; protein replacement therapy; crop;  
XX

KM	nutritional value; seed oil content; ds.
XX	
OS	Synthetic.
PN	
XN	MO200255719-A2.
XX	
PD	18-JUL-2002.
XX	
PF	09-JAN-2002; 2002MO-US01123.
XX	
PR	09-JAN-2001; 2001US-0758962.
XX	
PA	(BIOS-) BIOSOURCE GENETICS CORP.
PI	
DR	Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX	WPI; 2002-557829/59.
PT	
PT	New polypeptides contained in plant virus expression vectors, as gene
PT	expression tools, in protein replacement therapy or for intervening in
PT	a metabolic pathway to improve the nutritional value of a crop or alter
XX	the oil content of seeds
PS	
XX	Claim 4; Fig 4; 33pp; English.
XX	
CC	This invention describes a novel isolated polynucleotide comprising an
CC	internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC	frame (ORF) encoding a peptide and an ORF2 encoding a viral protein,
CC	where the IRES is located between ORF1 and ORF2. The novel
CC	polynucleotides of the invention are used in the construction of a
CC	recombinant potato virus X-based viral vector containing a nucleic acid
CC	construct comprising a bicistronic message with an intervening IRES. The
CC	constructs are used in a method for regulating the rate at which a virus
CC	infection spreads in a host. Regulation is achieved by placing the
CC	nucleic acid construct comprising an internal ribosomes entry site
CC	upstream of a coat protein gene, where the IRES is chosen by the rate of
CC	infection of the viral vector on a host in the presence of that IRES. The
CC	polynucleotide and vectors of the invention are useful for directing
CC	rapid and high-level expression of foreign genes in mature,
CC	differentiated, plant tissue. These are particularly useful in protein
CC	replacement therapy, or for intervening in a metabolic pathway to improve
CC	the nutritional value of a crop or alter the oil content of the seed.
CC	This sequence represents a fragment of the clone SC196 containing an
CC	IREScp region described in the method of the invention.
XX	
SQ	Sequence 180 BP; 60 A; 31 C; 52 G; 37 T; 0 other;
XX	
Query Match	23.4%; Score 44; DB 24; Length 180;
Best Local Similarity	100.0%; Pred. No. 9.3e-12;
Matches	44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	90 GGTGAGAAAAAGGGTGTAGTAGTAAGTATATAGTACAGACCGG 133
DB	89 GGTGAGAAAAAGGGTGTAGTAGTAAGTATATAGTACAGACCGG 132
XX	
RESULT 6	
AB076064	
ID	AB076064 standard; DNA; 181 BP.
XX	
AB076064;	
XX	
DT	30-SEP-2002 (first entry)
XX	
DE	Clone SC197 containing IREScp DNA sequence.
XX	
KM	IRES; internal ribosome entry site; viral proteoin; bicistronic;
KM	virus infection; coat protein; protein replacement therapy; crop;
XX	nutritional value; seed oil content; ds.
OS	Synthetic.
XX	
XN	MO200255719-A2.

XX	18-JUL-2002.
PD	
XX	09-JAN-2002; 2002MO-US01123.
XX	
PF	
XX	09-JAN-2001; 2001US-0758962.
PR	
XX	
XX	
PA	(BIOS-) BIOSOURCE GENETICS CORP.
PI	
PI	Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX	
DR	WPI; 2002-557829/59.
PT	
PT	New polypeptides contained in plant virus expression vectors, as gene
PT	expression tools, in protein replacement therapy or for intervening in
PT	a metabolic pathway to improve the nutritional value of a crop or alter
PT	the oil content of seeds
XX	
XX	
PS	Claim 4; Fig 4; 33pp; English.
XX	
CC	This invention describes a novel isolated polynucleotide comprising an
CC	internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC	frame (ORF) encoding a peptide and an ORF2 encoding a viral protein,
CC	where the IRES is located between ORF1 and ORF2. The novel
CC	polynucleotides of the invention are used in the construction of a
CC	recombinant potato virus X-based viral vector containing a nucleic acid
CC	construct comprising a bicistronic message with an intervening IRES. The
CC	constructs are used in a method for regulating the rate at which a virus
CC	infection spreads in a host. Regulation is achieved by placing the
CC	nucleic acid construct comprising an internal ribosomes entry site
CC	upstream of a coat protein gene, where the IRES is chosen by the rate of
CC	infection of the viral vector on a host in the presence of that IRES. The
CC	polynucleotide and vectors of the invention are useful for directing
CC	rapid and high-level expression of foreign genes in mature,
CC	differentiated, plant tissue. These are particularly useful in protein
CC	replacement therapy, or for intervening in a metabolic pathway to improve
CC	the nutritional value of a crop or alter the oil content of the seed.
CC	This sequence represents a fragment of the clone SC197 containing an
CC	IREScp region described in the method of the invention.
XX	
XX	
SQ	Sequence 181 BP; 61 A; 28 C; 53 G; 39 T; 0 other;
	Query Match 22.9%; Score 43; DB 24; Length 181;
	Best Local Similarity 100.0%; Pred. No. 2.8e-11;
	Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GTCGACGGTATCGATACCTTGATATCGAATCGTCGATTGG 43
DB	1 GTCGACGGTATCGATACCTTGATATCGAATCGTCGATTGG 43
RESULT 7	
ABQ76061	
ID	ABQ76061 standard; DNA; 122 BP.
XX	
AC	ABQ76061;
XX	
XX	
DT	30-SEP-2002 (first entry)
DE	
DE	Clone TXS-GFP-IRESmp-CP containing IRES DNA sequence.
XX	
XX	
KW	IRES; internal ribosome entry site; viral protein; bicistronic;
KW	virus infection; coat protein; protein replacement therapy; crop;
KW	nutritional value; seed oil content; ds.
OS	
XX	Synthetic.
XX	
XX	MO200255719-A2.
PN	
PD	18-JUL-2002.
XX	
XX	
XX	09-JAN-2002; 2002MO-US01123.
XX	





CC encoding the accessory factor TIP30-32.12.

XX Sequence 1353 BP; 412 A; 329 C; 294 G; 318 T; 0 other;

Query Match 18.6%; Score 35; DB 24; Length 1353;

Best Local Similarity 100.0%; Pred. No. 1.9e-07; Mismatches 0; Indels 0; Gaps 0;

DB 4 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 35  
 1 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 35  
 4 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 38

RESULT 12

AAV41999 standard; DNA; 2126 BP.

AC AAV41999;

XX 20-NOV-1998 (first entry)

DE Human GPR14 (G-protein coupled receptor polypeptide) gene.

KW G-protein coupled receptor polypeptide; GPR14; human; fungal infection;  
 KW bacterial infection; protozoan infection; viral infection; agonist;  
 KW antagonist; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 694..1863

FT /\*tag= a

FT /product= "human GPR14"

XX EP859052-A1.

PD 19-AUG-1998.

PF 18-NOV-1997; 97EP-0309252.

PR 27-JAN-1997; 97US-0789354.

PA (SMIK ) SMITHKLINE BEECHAM CORP.

PI Bergsma DJ, Shabon U;

DR WPI: 1998-429680/37.

DR P-PSDB; AAW59908.

PT New DNA sequence encoding G-protein coupled receptor polypeptide  
 PT GPR14 - and corresponding polypeptide, agonists, antagonists, etc.  
 XX Claim 3; Fig 1; 23pp; English.

CC This is the nucleotide sequence of the gene encoding the novel G-protein  
 CC coupled receptor polypeptide (GPR14), used in the method of the  
 CC invention. Human GPR14 polypeptides and polynucleotides can be used in  
 CC the treatment of infections such as bacterial, fungal, protozoan and  
 CC viral infections. Agonists and antagonists can be used to treat  
 CC conditions associated with Human GPR14 imbalance.

XX Sequence 2126 BP; 317 A; 763 C; 669 G; 377 T; 0 other;

Query Match 18.6%; Score 35; DB 19; Length 2126;

Best Local Similarity 100.0%; Pred. No. 1.8e-07; Mismatches 0; Indels 0; Gaps 0;

DB 1 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 35  
 1 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 35  
 30 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 64

RESULT 13

AAZ11174 standard; CDNA; 2126 BP.

XX AAZ11174;

DT 04-NOV-1999 (first entry)

DE Human GPR14 coding sequence.

KW GPR14; G-protein-coupled receptor; ischemic coronary artery disease;  
 KW atherosclerosis; metabolic disease; CHF/myocardial dysfunction; migraine;  
 KW arrhythmia; restenosis; hypertension; hypotension; pulmonary disease;  
 KW fibrotic vasculopathy; cerebrovascular event; neurogenic inflammation;  
 KW haematopoietic disorder; adult respiratory distress syndrome; ARDS;  
 KW cancer; autoimmune disease; therapy; human; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 694..1863

FT /\*tag= a

FT /product= GPR14

PD 12-AUG-1999.

PF 27-JAN-1999; 99WO-US01634.

PR 15-JAN-1999; 99US-0232857.

PR 09-FEB-1998; 98US-0074075.

PR 10-APR-1998; 98US-0058725.

PA (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM LAB PHARM.

PI Alvar NV, Al-Barazanji K, Ames RS, Arnold AR, Bergsma DJ;

PI Chambers J, Douglas SA, Foley JJ, Gout B, Khandoudi N;

PI Sarau HM, Shabon U, Willette RN;

DR WPI: 1999-527305/44.

DR P-PSDB; AAY32920.

XX Human G-protein-coupled receptor GPR14, useful for identifying

XX agonists and antagonists

XX Claim 4; Fig 1; 64pp; English.

CC This sequence encodes the human G-protein-coupled receptor, GPR14, of the  
 CC invention. Human GPR14, polypeptide and polynucleotides are useful in  
 CC methods for treatment of, e.g. ischemic coronary artery disease,  
 CC atherosclerosis, metabolic diseases, CHF/myocardial dysfunction,  
 CC arrhythmias, restenosis, hypertension and hypotension, pulmonary disease,  
 CC fibrotic vasculopathies, cerebrovascular events, neurogenic  
 CC inflammation/migraine, haematopoietic disorders, adult respiratory  
 CC distress syndrome (ARDS), cancer, autoimmune diseases, etc. The methods  
 CC can be used to identify agonists and antagonists of human and rat GPR14.

XX Sequence 2126 BP; 317 A; 763 C; 669 G; 377 T; 0 other;

Query Match 18.6%; Score 35; DB 20; Length 2126;

Best Local Similarity 100.0%; Pred. No. 1.8e-07; Mismatches 0; Indels 0; Gaps 0;

DB 1 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 35  
 1 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 35  
 30 GTCGACGGTATCGATTAAGCTTGATATCGAATTGCT 64

RESULT 14

AAK9298 standard; CDNA; 3236 BP.



```

XX AC AAK9298;
XX XX
XX 11-JUN-2002 (first entry)
XX DE cDNA of the human cancer suppressor gene 98.
XX XX
XX CYTOSTATIC; anti-HIV; human; cancer suppressor gene 98; antagonist;
XX KW DNA recombination; cancer; HIV infection; gene; ss.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 76..2760
XX FT /*tag- a
XX FT /product- "Protein of human cancer suppressor gene 98"
XX PN CN1328030-A.
XX XX
XX PD 26-DEC-2001.
XX PF 12-JUN-2000; 2000CN-0116437.
XX PR 12-JUN-2000; 2000CN-0116437.
XX XX
XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX XX
XX P1 Mao Y, Xie Y;
XX XX
XX DR WPI: 2002-258228/31.
XX DR P-PSDB; AAO20441.
XX XX
XX PT Polypeptide-human cancer suppressor gene 98, useful for treating
XX PT diseases such as cancer and HIV infection.
XX PS Claim 6; Page 25-27 Disclosure; 37pp; Chinese.
XX XX
XX CC The invention relates to a novel protein of the human cancer suppressor
XX CC gene 98, the encoding polynucleotide, a process for preparing the
XX CC polynucleotide by DNA recombination, and an antagonist to the
XX CC polynucleotide. The protein is useful for treating diseases such as
XX CC cancer, HIV infection, etc. This polynucleotide sequence represents the
XX CC cDNA of the human cancer suppressor gene 98 of the invention.
XX XX
XX SQ Sequence 3236 BP; 892 A; 638 C; 745 G; 961 T; 0 other;

Query Match 18.6%; Score 35; DB 24; Length 3236;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGGTATCGATAAGCTTGATATCGAATTGCT 35
    |||||||
DB 31 GTCGACGGTATCGATAAGCTTGATATCGAATTGCT 65

RESULT 15
ABN83030
ID ABN83030 standard; cDNA; 3905 BP.
XX
XX AC ABN83030;
XX XX
XX DT 14-AUG-2002 (first entry)
XX DE Human X2 theca repressor 10.34 cDNA.
XX XX
XX KM Human; X2 theca repressor 10.34; repressor; immunopathy; gene; ss.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 3574..3858
XX FT /*tag- a
XX FT /product- "X2 theca repressor 10.34"

```

```

XX PN CN1331139-A.
XX XX
XX PD 16-JAN-2002.
XX PF 26-JUN-2000; 2000CN-0116749.
XX PR 26-JUN-2000; 2000CN-0116749.
XX XX
XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX XX
XX P1 Mao Y, Xie Y;
XX XX
XX DR WPI: 2002-292849/34.
XX DR P-PSDB; ABB81691.
XX XX
XX PT Polypeptide-human X2 theca repressor 10.34 and polynucleotide for
XX PT coding it -
XX PS Claim 6; Page 25-27; 34pp; Chinese.
XX XX
XX CC The sequence encodes the novel human X2 theca repressor 10.34 polypeptide
XX CC of the invention. The polypeptide is useful for treating diseases such as
XX CC immunopathy. The antagonist of the polypeptide and its medical action.
XX CC and the application of the polynucleotide are also disclosed.
XX XX
XX SQ Sequence 3905 BP; 1245 A; 731 C; 701 G; 1228 T; 0 other;

Query Match 18.6%; Score 35; DB 24; Length 3905;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGGTATCGATAAGCTTGATATCGAATTGCT 35
    |||||||
DB 31 GTCGACGGTATCGATAAGCTTGATATCGAATTGCT 65

Search completed: February 19, 2003, 07:13:20
Job time : 231 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2003, 07:08:12 ; Search time 45 Seconds  
(without alignments)  
1281.227 Million cell updates/sec

Title: US-09-758-962-1

Perfect score: 188  
Sequence: 1 gtcgacgcatcgataagct.....gaagaagagctcaccatcg 188

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 segs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A-COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B-COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/5A-COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/5B-COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS-COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	18.6	2126	2 US-08-789-354-1	Sequence 1, Appl
2	35	18.6	2126	3 US-09-110-937-1	Sequence 1, Appl
3	35	18.6	2126	3 US-09-058-725B-1	Sequence 1, Appl
4	35	18.6	2126	3 US-09-232-857-1	Sequence 1, Appl
5	35	18.6	7560	4 US-09-103-478-4	Sequence 4, Appl
6	35	18.6	7560	4 US-09-193-931C-4	Sequence 4, Appl
7	34	18.1	924	3 US-08-983-409-3	Sequence 3, Appl
8	34	18.1	1270	5 PCT-US95-10403-1	Sequence 1, Appl
9	34	18.1	1318	3 US-09-183-253-3	Sequence 1, Appl
10	34	18.1	1471	4 US-09-492-985-11	Sequence 11, Appl
11	34	18.1	1638	2 US-08-838-219B-8	Sequence 8, Appl
12	34	18.1	1638	3 US-09-233-336A-8	Sequence 8, Appl
13	34	18.1	1638	3 US-09-233-752A-8	Sequence 8, Appl
14	34	18.1	1638	4 US-09-402-036-8	Sequence 8, Appl
15	34	18.1	1638	4 US-09-904-226-8	Sequence 8, Appl
16	34	18.1	1679	3 US-08-676-882-1	Sequence 3, Appl
17	34	18.1	1703	5 PCT-US95-10403-3	Sequence 3, Appl
18	34	18.1	1958	4 US-09-215-221-9	Sequence 15, App
19	34	18.1	2031	2 US-08-467-948A-3	Sequence 3, Appl
20	34	18.1	2185	3 US-08-467-947A-3	Sequence 3, Appl
21	34	18.1	2200	5 US-08-272-255-21	Sequence 21, Appl
22	34	18.1	2200	5 PCT-US95-08565-21	Sequence 21, Appl
23	34	18.1	2605	2 US-08-680-395-13	Sequence 4, Appl
24	34	18.1	3344	1 US-07-718-575-13	Sequence 13, Appl
25	34	18.1	3344	1 US-08-481-206-13	Sequence 13, Appl
26	34	18.1	3344	2 US-08-486-269A-13	Sequence 13, Appl
27	34	18.1	3344	2 US-08-486-269A-13	Sequence 13, Appl

28	34	18.1	5470	2 US-08-443-861-1	Sequence 1, Appl
29	34	18.1	5470	4 US-08-193-829B-1	Sequence 1, Appl
30	33	17.6	57	4 US-09-203-681-4	Sequence 4, Appl
31	33	17.6	78	4 US-09-462-645C-25	Sequence 25, Appl
32	33	17.6	84	4 US-09-462-645C-26	Sequence 26, Appl
33	33	17.6	88	1 US-08-144-602B-15	Sequence 15, Appl
34	33	17.6	129	1 US-08-752-238-4	Sequence 4, Appl
35	33	17.6	129	3 US-09-085-603B-4	Sequence 21, Appl
36	33	17.6	129	3 US-09-085-603B-21	Sequence 22, Appl
37	33	17.6	129	3 US-09-323-872A-26	Sequence 39, Appl
38	33	17.6	147	4 US-09-116-492A-39	Sequence 14, Appl
39	33	17.6	229	4 US-08-387-707-14	Sequence 14, Appl
40	33	17.6	332	4 US-08-405-271A-14	Sequence 306, App
41	33	17.6	332	4 US-09-643-597-306	Sequence 4, Appl
42	33	17.6	457	3 US-08-928-799A-4	Sequence 10, Appl
43	33	17.6	516	3 US-08-991-789A-10	Sequence 10, Appl
44	33	17.6	545	4 US-09-062-451-10	Sequence 10, Appl
45	33	17.6	545	4 US-09-062-451-10	Sequence 10, Appl

#### ALIGNMENTS

RESULT 1  
US-08-789-354-1  
Sequence 1, Application US/08789354

Patent No. 5851798

GENERAL INFORMATION:

APPLICANT: Shadon, Usman

APPLICANT: Bergsma, Derek

TITLE OF INVENTION: Cloning of Human GPR14 Re

NUMBER OF INVENTIONS: 2

CORRESPONDENCE ADDRESS:

ADDRESS: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Fastseq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/789,354

FILING DATE: 27-JAN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Han, William T

REGISTRATION NUMBER: 34,344

REFERENCE/DOCKET NUMBER: P50610

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5219

TELEFAX: 610-270-4026

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2126 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-789-354-1

Query Match 18.6% Score 35; DB 2; Length 2126;

Best Local Similarity 100.0%; Pred. No. 2e-08;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACGGTATCGATTAAGCTTGATCGAATTCGT 35  
|||||  
Db 30 GTGACGGTATCGATTAAGCTTGATCGAATTCGT 64

## RESULT 2

US-09-110-937-1  
; Sequence 1, Application US/09110937A  
; Patent No. 6005074  
; GENERAL INFORMATION:  
; APPLICANT: SHABON, USMAN  
; APPLICANT: BERGSMAN, DEREK  
; TITLE OF INVENTION: CLONING OF HUMAN GPR14 RECEPTOR  
; FILE REFERENCE: PS0610-1  
; CURRENT APPLICATION NUMBER: US/09/110,937A  
; CURRENT FILING DATE: 1998-07-06  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO: 1  
; LENGTH: 2126  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
US-09-110-937-1

Query Match 18.6%; Score 35; DB 3; Length 2126;  
Best Local Similarity 100.0%; Pred. No. 2e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACGGTATCGATTAAGCTTGATCGAATTCGT 35  
|||||  
Db 30 GTGACGGTATCGATTAAGCTTGATCGAATTCGT 64

## RESULT 3

US-09-058-7258-1  
; Sequence 1, Application US/090587258  
; Patent No. 6133420  
; GENERAL INFORMATION:

APPLICANT: Ames, Robert  
APPLICANT: Sarau, Henry  
APPLICANT: Foley, James  
APPLICANT: Chamber, Jon  
TITLE OF INVENTION: A Method of Finding Angonist  
TITLE OF INVENTION: and Antagonist To Human and Rat GPR 14  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESS: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/058,7258  
FILING DATE: April 10, 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/789,354  
FILING DATE: 27-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Han, William T.  
REGISTRATION NUMBER: 34,344  
REFERENCE/DOCKET NUMBER: GP50005-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5219  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2126 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-058-7258-1

Query Match 18.6%; Score 35; DB 3; Length 2126;  
Best Local Similarity 100.0%; Pred. No. 2e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACGGTATCGATTAAGCTTGATCGAATTCGT 35  
|||||  
Db 30 GTGACGGTATCGATTAAGCTTGATCGAATTCGT 64

## RESULT 4

US-09-232-857-1  
; Sequence 1, Application US/09232857  
; Patent No. 6159700  
; GENERAL INFORMATION:

APPLICANT: DOUGLAS, STEPHEN  
APPLICANT: WILLETT, ROBERT  
APPLICANT: AIYAR, NAMBI  
APPLICANT: ROMANIC, ANNE  
APPLICANT: KHANDOUDI, NASSIRAH  
APPLICANT: GOUT, BERNARD  
APPLICANT: AL-BARAZANJI, KAMAL  
APPLICANT: AMES, ROBERT S.  
APPLICANT: FOLEY, JAMES J.  
APPLICANT: SARAU, HENRY  
APPLICANT: CHAMBERS, JON K.  
APPLICANT: SHABON, USMAN  
APPLICANT: BERGSMAN, DEREK  
TITLE OF INVENTION: A METHOD OF FINDING AGONIST  
TITLE OF INVENTION: AND ANTAGONIST TO HUMAN AND RAT GPR14  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESS: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/232,857  
FILING DATE: 15-JAN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/789,354  
FILING DATE: 27-JAN-1997  
APPLICATION NUMBER: 60/074,075  
FILING DATE: 09-FEB-1998  
APPLICATION NUMBER: 08/058,725  
FILING DATE: 10-APR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F.  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GP-50005-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2126 base pairs  
TYPE: nucleic acid



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/983,409  
FILING DATE: 20-JAN-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB96/01894  
FILING DATE: 06-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9516961.1  
FILING DATE: 18-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hohenschultz, Lisa D.  
REGISTRATION NUMBER: 33,712  
REFERENCE/DOCKET NUMBER: SEE45001/UST  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 886-1699  
INFORMATION FOR SEQ. ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 924 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-983-409-3

Query Match 18.1%; Score 34; DB 3; Length 924;  
Best Local Similarity 100.0%; Pred. No. 6.6e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGGTATCGATTAAGCTTGATTCGAATTGC 34  
DB 12 GTGACGGTATCGATTAAGCTTGATTCGAATTGC 45

## RESULT 8

PCT-US95-10403-1  
Sequence 1, Application PC/TUS9510403  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation  
TITLE OF INVENTION: GA4 DNA, Protein and Methods of Use  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein and Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10403  
FILING DATE: 15 August 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/291,939  
FILING DATE: 16 August 1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbale, Michele A.  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 0609.408PC00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1270 base pairs

TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1183  
PCT-US95-10403-1

Query Match 18.1%; Score 34; DB 5; Length 1270;  
Best Local Similarity 100.0%; Pred. No. 6.4e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGGTATCGATTAAGCTTGATTCGAATTGC 34  
DB 6 GTGACGGTATCGATTAAGCTTGATTCGAATTGC 39

## RESULT 9

US-09-183-253-3/c  
Sequence 3, Application US/09183253  
Patent No. 6043054  
GENERAL INFORMATION:  
APPLICANT: VAMTER, LISA  
APPLICANT: STAMMER, MELANIE  
TITLE OF INVENTION: NOVEL COMPOUNDS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/183,253  
FILING DATE: 30-OCT-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9817907.0  
FILING DATE: 17-AUG-1998  
APPLICATION NUMBER: 60/075,306  
FILING DATE: 20-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F.  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GP-70395  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TEXT: 846169  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-183-253-3

Query Match 18.1%; Score 34; DB 3; Length 1318;  
Best Local Similarity 100.0%; Pred. No. 6.4e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGGTATCGATTAAGCTTGATTCGAATTGC 34  
DB 1318 GTGACGGTATCGATTAAGCTTGATTCGAATTGC 1285

## RESULT 10

US-09-492-985-11/C  
Sequence 11, Application US/09492985  
Patent No. 6376240  
GENERAL INFORMATION:  
APPLICANT: Song, An M.  
APPLICANT: Chen, Ya-Fen  
APPLICANT: Krensky, Alan M.  
TITLE OF INVENTION: RFLAT-1: A Transcription Factor That  
FILE REFERENCE: SUN-113P  
CURRENT APPLICATION NUMBER: US/09/492,985  
CURRENT FILING DATE: 2000-01-27  
EARLIER APPLICATION NUMBER: 60/117,576  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0.  
SEQ ID NO 11  
LENGTH: 1471  
TYPE: DNA  
ORGANISM: mouse  
US-09-492-985-11

Query Match 18.1%; Score 34; DB 4; Length 1471;  
Best Local Similarity 100.0%; Pred. No. 6.4e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1443 GTCACGCTATCGATAGCTTGATTCG 34  
1443 GTCACGCTATCGATAGCTTGATTCG 1410

## RESULT 11

US-08-838-219B-8/C  
Sequence 8, Application US/08838219B  
Patent No. 5877012

GENERAL INFORMATION:  
APPLICANT: Warren, Gregory M  
APPLICANT: Koziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalin M  
APPLICANT: Kostichka, N. Kristy  
APPLICANT: Duck, Nicholas B  
APPLICANT: Estruch, Juan J  
TITLE OF INVENTION: A No. 5877012el Class of Proteins for the  
TITLE OF INVENTION: Control of Plant Pests  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,219B  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/463,483  
FILING DATE: 06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/314,594  
FILING DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,018

## FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/037,057  
FILING DATE: 23-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40,403  
REFERENCE/DOCKET NUMBER: CGC 1925  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 519-541-8582  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1638 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..1191  
OTHER INFORMATION:  
OTHER INFORMATION: encoding VIP3a(a) receptor"

Query Match 18.1%; Score 34; DB 2; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1633 GTCACGCTATCGATAGCTTGATTCG 34  
1633 GTCACGCTATCGATAGCTTGATTCG 1600

## RESULT 12

US-09-233-336A-8/C  
Sequence 8, Application US/09233336A  
Patent No. 6107279

GENERAL INFORMATION:  
APPLICANT: Warren, Gregory M  
APPLICANT: Koziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalin M  
APPLICANT: Kostichka, N. Kristy  
APPLICANT: Duck, Nicholas B  
APPLICANT: Estruch, Juan J  
TITLE OF INVENTION: A No. 6107279el Class of Proteins for the  
TITLE OF INVENTION: Control of Plant Pests  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/233,336A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/838,219  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/314,594

FILED DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,018  
FILING DATE: 23-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/037,057  
FILING DATE: 25-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40,403  
REFERENCE/DOCKET NUMBER: CGC 1925  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8582  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1638 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..1191  
OTHER INFORMATION: /Product-"translation of cDNA  
US-09-233-336A-8  
OTHER INFORMATION: encoding VIP3A(a) receptor"

Query Match 18.1%; Score 34; DB 3; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGGTATCGATTAAGCTTGATTCGAATTCG 34  
DB 1633 GTGACGGTATCGATTAAGCTTGATTCGAATTCG 1600

RESULT 13  
US-09-233-752A-8/c  
Sequence 8, Application US/09233752A  
Patent No. 6137033  
GENERAL INFORMATION:  
APPLICANT: Warren, Gregory W  
APPLICANT: Kozziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalin M  
APPLICANT: Kostichka, N. Kristy  
APPLICANT: Duck, Nicholas B  
APPLICANT: Estruch, Juan J  
TITLE OF INVENTION: A No. 6137033el Class of Proteins for the  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/233,752A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/838,219

FILED DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/314,594  
FILING DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,018  
FILING DATE: 23-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40,403  
REFERENCE/DOCKET NUMBER: CGC 1925  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8582  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1638 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..1191  
OTHER INFORMATION: /Product-"translation of cDNA  
US-09-233-752A-8  
OTHER INFORMATION: encoding VIP3A(a) receptor"

Query Match 18.1%; Score 34; DB 3; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGGTATCGATTAAGCTTGATTCGAATTCG 34  
DB 1633 GTGACGGTATCGATTAAGCTTGATTCGAATTCG 1600

RESULT 14  
US-09-402-036-8/c  
Sequence 8, Application US/09402036  
Patent No. 6291156  
GENERAL INFORMATION:  
APPLICANT: Estruch, Juan J.  
APPLICANT: Yu, Cao-Guo  
APPLICANT: Warren, Gregory W.  
APPLICANT: Desai, Nalin M  
APPLICANT: Kozziel, Michael  
APPLICANT: Nye, Gordon  
TITLE OF INVENTION: Plant Pest Control  
FILE REFERENCE: S-21284C  
CURRENT APPLICATION NUMBER: US/09/402,036  
CURRENT FILING DATE: 2000-02-08  
PRIOR APPLICATION NUMBER: PCT/EP98/01952  
PRIOR FILING DATE: 1998-04-02  
PRIOR APPLICATION NUMBER: 08/838,219  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: 08/832,263  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: 08/832,265  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: 08/463,483  
PRIOR FILING DATE: 1995-06-05  
PRIOR APPLICATION NUMBER: 08/314,594  
PRIOR FILING DATE: 1994-09-09  
PRIOR APPLICATION NUMBER: 08/218,018  
PRIOR FILING DATE: 1994-03-24  
PRIOR APPLICATION NUMBER: 08/037,057  
PRIOR FILING DATE: 1993-03-25  
NUMBER OF SEQ ID NOS: 22

Search completed: February 19, 2003, 08:30:42  
Job time : 50 secs

SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1638  
TYPE: DNA  
ORGANISM: Agrotis Ipsilon  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)..(1189)  
OTHER INFORMATION: Translation of cDNA encoding VIP3A(a) receptor  
OTHER INFORMATION: from Black cutworm  
US-09-402-036-8

Query Match 18.1%; Score 34; DB 4; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGGTATCGATAGCTGATATCGAATTGC 34  
|||||  
Db 1633 GTCGACGGTATCGATAGCTGATATCGAATTGC 1600

## RESULT 15

US-09-904-226-8/c  
Sequence 8, Application US/09904226  
Patent No. 6429360  
GENERAL INFORMATION:  
APPLICANT: Estruch, Juan J.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Desai, Nalini  
APPLICANT: Koziel, Michael  
APPLICANT: Nye, Gordon  
TITLE OF INVENTION: Plant Pest Control  
FILE REFERENCE: S-21284D  
CURRENT APPLICATION NUMBER: US/09/904,226  
CURRENT FILING DATE: 2001-07-12  
PRIOR APPLICATION NUMBER: PCT/EP98/01952  
PRIOR FILING DATE: 1998-04-02  
PRIOR APPLICATION NUMBER: 08/838,219  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: 08/832,263  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: 08/832,265  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: 08/463,483  
PRIOR FILING DATE: 1995-06-05  
PRIOR APPLICATION NUMBER: 08/314,594  
PRIOR FILING DATE: 1994-09-09  
PRIOR APPLICATION NUMBER: 08/218,018  
PRIOR FILING DATE: 1994-03-24  
PRIOR APPLICATION NUMBER: 08/037,057  
PRIOR FILING DATE: 1993-03-25  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1638  
TYPE: DNA  
ORGANISM: Agrotis Ipsilon  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)..(1189)  
OTHER INFORMATION: Translation of cDNA encoding VIP3A(a) receptor  
OTHER INFORMATION: from Black cutworm  
US-09-904-226-8

Query Match 18.1%; Score 34; DB 4; Length 1638;  
Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACGGTATCGATAGCTGATATCGAATTGC 34  
|||||  
Db 1633 GTCGACGGTATCGATAGCTGATATCGAATTGC 1600



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2003, 07:00:11 ; Search time 1949 seconds  
(without alignments)  
1562.212 Million cell updates/sec

Title: US-09-758-962-1

Perfect score: 188  
Sequence: 1 gtcgacggtatcgtataagct.....gaagaagaagctccaccatg 188

Scoring table: OLIGO NDC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:  
1: em\_estba:\*  
2: em\_esthm:\*  
3: em\_estln:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_num:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	38	20.2	304	9	AU167182 AU167182
2	36	19.1	818	13	BM141055 BM141055
3	35	18.6	495	13	BM139700 BM139700
4	35	18.6	919	13	BM351825 BM351825
5	35	18.6	945	13	BM351824 BM351824
6	35	18.6	1300	13	BM351823 BM351823

C	7	34	18.1	52	10	AM172124	AM172124
C	8	34	18.1	53	10	AM042087	AM042087
C	9	34	18.1	53	10	AM626527	AM626527
C	10	34	18.1	54	10	AM172172	AM172172
C	11	34	18.1	54	12	BE758345	BE758345
C	12	34	18.1	66	10	AM626545	AM626545
C	13	34	18.1	68	10	AM159927	AM159927
C	14	34	18.1	91	9	AI066878	AI066878
C	15	34	18.1	92	9	AA680557	AA680557
C	16	34	18.1	92	14	BO821730	BO821730
C	17	34	18.1	105	14	BO819685	BO819685
C	18	34	18.1	143	9	AL837413	AL837413
C	19	34	18.1	175	9	AL835751	AL835751
C	20	34	18.1	181	12	BG360460	BG360460
C	21	34	18.1	186	10	AM740238	AM740238
C	22	34	18.1	190	10	AM739962	AM739962
C	23	34	18.1	206	10	AM743048	AM743048
C	24	34	18.1	212	10	AM739871	AM739871
C	25	34	18.1	215	10	AM739826	AM739826
C	26	34	18.1	217	10	AM740190	AM740190
C	27	34	18.1	217	10	AM740458	AM740458
C	28	34	18.1	217	10	BE530967	BE530967
C	29	34	18.1	218	9	AJ407770	AJ407770
C	30	34	18.1	223	10	AM739964	AM739964
C	31	34	18.1	227	10	AM742899	AM742899
C	32	34	18.1	231	10	AM739775	AM739775
C	33	34	18.1	232	10	AM740417	AM740417
C	34	34	18.1	233	10	AM740289	AM740289
C	35	34	18.1	240	9	AJ477003	AJ477003
C	36	34	18.1	241	10	AM739737	AM739737
C	37	34	18.1	253	10	AM740254	AM740254
C	38	34	18.1	254	10	AM740521	AM740521
C	39	34	18.1	258	9	AI052896	AI052896
C	40	34	18.1	259	10	AM739879	AM739879
C	41	34	18.1	259	10	AM740178	AM740178
C	42	34	18.1	259	10	AM740378	AM740378
C	43	34	18.1	263	14	BO815505	BO815505
C	44	34	18.1	267	10	AM739838	AM739838
C	45	34	18.1	274	13	BI119516	BI119516

## ALIGNMENTS

RESULT 1  
AU167182/c  
LOCUS  
DEFINITION  
AU167182 01-br-ad cDNA Oryzias latipes cDNA clone br0230, mRNA  
ACCESSION  
AU167182  
VERSION  
AU167182.1 GI:12589251  
KEYWORDS  
SOURCE  
ORGANISM  
Japanese medaka.  
Oryzias latipes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
Belontiiformes; Adiantichthyidae; Oryziinae; Oryzias.  
1 (bases 1 to 304)  
Mita, K., Ishikawa, Y. and Yamachi, M.  
Establishment of cDNA database of medaka, Oryzias latipes  
Unpublished (2001)  
CONTACT: Mita K  
Genome Research Group  
National Institute of Radiological Sciences  
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
Email: kmits@ris.go.jp  
method: uni-directional sequence direction: sequenced from T3 primer  
(5' -> 3').

FEATURES  
SOURCE  
1. 304  
/organism="Oryzias latipes"  
/strain="HNI"

Microbiology and Immunology  
Stanford University  
D-305 Fairchild Building, 300 Pasteur Drive Stanford, CA 94305, USA  
Tel: 650 723 7266  
Fax: 650 723 6853  
Email: usingsh@stanford.edu  
Seq primer: 77  
High quality sequence start: 20  
High quality sequence stop: 700

FEATURES  
Source

POLYA=yes. Location/Qualifiers

1. 919

/organism="Toxoplasma gondii"

/strain="ME49"

/db\_xref="taxon:5811"

/clone="Tgzz44h01.s1"

/clone.lib="TgME49 Invivo Bradyzoite cDNA size selected"

/dev\_stage="Bradyzoite"

/lab\_host="DH10"

/note="Vector: Bluescript II SK-; Site\_1: EcoRI; Site\_2: NotI; Mature bradyzoites were obtained from infected mouse brains by percoll density centrifugation. The original library was constructed by Steve Parmley, Palo Alto Medical Foundation. cDNAs were synthesized by priming with oligo d(T) and directionally cloned into the EcoRI/NotI sites of lambda gt11. Warning: the library contains a small percentage of host cDNAs derived from mouse cells. Inserts from this cDNA library were excised with NotI and EcoRI, size selected in a range of 0.7 - 2.0 kb and subcloned into Bluescript II SK- (Adrian Hehl, Ian Manger and John Boothroyd, Stanford University)"

BASE COUNT 217 a 219 c 232 g 231 t 20 others

Query Match 18.6%; Score 35; DB 13; Length 919; Best Local Similarity 100.0%; Pred. No. 8.6e-07;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGACGGTATCGATAGCTTGATATCGAATTCGT 35  
|||||  
Db 30 GTGACGGTATCGATAGCTTGATATCGAATTCGT 64

RESULT 5

BM351824

LOCUS BM351824 945 bp mRNA linear EST 29-JUL-2002

DEFINITION Tgzz40h11.s1 TgME49 Invivo Bradyzoite cDNA size selected Toxoplasma

BM351824

ACCESSION BM351824.1 GI:18084182

VERSION EST.

KEYWORDS Toxoplasma gondii.

SOURCE Toxoplasma gondii.

ORGANISM Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Journal Contact: Upinder Singh

Microbiology and Immunology

D-305 Fairchild Building, 300 Pasteur Drive Stanford, CA 94305, USA

Tel: 650 723 7296

Fax: 650 723 6853

Email: usin@stanford.edu

Seq primer: 77

High quality sequence start: 20

High quality sequence stop: 700

POLYA-No.

Location/Qualifiers

1. 945

/organism="Toxoplasma gondii"

/strain="ME49"

/db\_xref="taxon:5811"

/clone="Tgzz44h01.s1"

/clone.lib="TgME49 Invivo Bradyzoite cDNA size selected"

/dev\_stage="Bradyzoite"

/lab\_host="DH10"

/note="Vector: Bluescript II SK-; Site\_1: EcoRI; Site\_2:

BASE COUNT 199 a 182 c 266 g 261 t 37 others

Query Match 18.6%; Score 35; DB 13; Length 945; Best Local Similarity 100.0%; Pred. No. 8.6e-07;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGACGGTATCGATAGCTTGATATCGAATTCGT 35  
|||||  
Db 46 GTGACGGTATCGATAGCTTGATATCGAATTCGT 80

RESULT 6 1300 bp mRNA linear EST 29-JUL-2002

LOCUS BM351823

DEFINITION Tgzz40d09.s1 TgME49 Invivo Bradyzoite cDNA size selected Toxoplasma

BM351823

ACCESSION BM351823.1 GI:18084181

VERSION EST.

KEYWORDS Toxoplasma gondii.

SOURCE Toxoplasma gondii.

ORGANISM Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Journal Contact: Upinder Singh

Microbiology and Immunology

D-305 Fairchild Building, 300 Pasteur Drive Stanford, CA 94305, USA

Tel: 650 723 7296

Fax: 650 723 6853

Email: usin@stanford.edu

Seq primer: 77

High quality sequence start: 50

High quality sequence stop: 750

POLYA-No.

Location/Qualifiers

1. 1300

/organism="Toxoplasma gondii"

/strain="ME49"

/db\_xref="taxon:5811"

/clone="Tgzz40d09.s1"

/clone.lib="TgME49 Invivo Bradyzoite cDNA size selected"

/dev\_stage="Bradyzoite"

/lab\_host="DH10"

/note="Vector: Bluescript II SK-; Site\_1: EcoRI; Site\_2: NotI; Mature bradyzoites were obtained from infected mouse brains by percoll density centrifugation. The original library was constructed by Steve Parmley, Palo Alto Medical Foundation. cDNAs were synthesized by priming with oligo d(T) and directionally cloned into the EcoRI/NotI sites of lambda gt11. Warning: the library contains a small percentage of host cDNAs derived from mouse cells. Inserts from this cDNA library were excised with NotI and EcoRI, size selected in a range of 0.7 - 2.0 kb and subcloned into Bluescript II SK- (Adrian Hehl, Ian Manger and John Boothroyd, Stanford University)"

```

BASE COUNT      284 a      296 c      316 g      290 t      114 others
ORIGIN
Query Match      18.6%; Score 35; DB 13; Length 1300;
Best Local Similarity 100.0%; Pred. No. 8.6e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GTCGACGGTATCGATAGCTTGATATCGAATTCGT 35
        |||||||
Db      56 GTCGACGGTATCGATAGCTTGATATCGAATTCGT 90

RESULT 7
LOCUS      AM172124      52 bp      mRNA      linear      EST 15-NOV-1999
DEFINITION SM4CAK07B01SK Brugia malayi L4 cDNA (SAM99MLW-Bml4) Brugia malayi
VERSION      AM172124
KEYWORDS
SOURCE      Brugia malayi.
ORGANISM      Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioides;
              Onchocercidae; Brugia.
REFERENCE      1 (bases 1 to 52)
AUTHORS      Williams, S. A.
TITLE      Genes expressed in fourth stage larvae of Brugia malayi
JOURNAL      Unpublished (1999)
COMMENT      Contact: Steven A. Williams
              Molecular Parasitology
              Smith College Department of Biological Sciences
              Department of Biological Sciences, Clark Science Center, Smith
              College, Northampton, MA, 01063, USA
              Tel: 4135853826
              Fax: 4135853786
              Email: genomesmith.edu
              Seq primer: pluescript SK.
              Location/Qualifiers
                1..52
                /organism="Brugia malayi"
                /db_xref="taxon:6279"
                /clone="SM4CAK07B01"
                /clone_1lb="Brugia malayi L4 cDNA (SAM99MLW-Bml4)"
                /dev_stage="larval stage four"
                /lab_host="XLI-Blue MRF"
                /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                Xho I; Lymphatic filarial nematode parasite of humans.
                mRNA was prepared from L4s isolated from the peritoneal
                cavity of jirds and converted to double-stranded cDNA
                using reverse transcriptase and oligo(dT) followed by
                RNase H and DNA pol I. The library has 2.7 x 105
                independent recombinants and the average insert size is
                approx. 1050bp. The library was constructed by Michelle
                Lizotte-Waniewski. The library is available from Dr. S.A.
                Williams, email:genomesmith.edu."

BASE COUNT      11 a      15 c      11 g      15 t
ORIGIN
Query Match      18.1%; Score 34; DB 10; Length 52;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GTCGACGGTATCGATAGCTTGATATCGAATTCG 34
        |||||||
Db      43 GTCGACGGTATCGATAGCTTGATATCGAATTCG 10

RESULT 9
LOCUS      AM626527      53 bp      mRNA      linear      EST 30-MAR-2000
DEFINITION SMOVL3CAN65A07SK Onchocerca volvulus infective larva cDNA
              (SAM99MLW-OVL3) Onchocerca volvulus cDNA clone SMOVL3CAN65A07 5',
              mRNA sequence.
ACCESSION      AM626527
VERSION      AM626527.1 GI:7342392
KEYWORDS
SOURCE      Onchocerca volvulus.
ORGANISM      Onchocerca volvulus.
              Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioides;
              Onchocercidae; Onchocerca.
REFERENCE      1 (bases 1 to 53)
AUTHORS      Williams, S.A., Lu, W., Lizotte-Waniewski, M. and Laney, S.J.
TITLE      Genes expressed in infective third stage larvae of Onchocerca
              volvulus
JOURNAL      Unpublished (1995)
COMMENT      Contact: Steven A. Williams
              Molecular Parasitology
              Smith College Department of Biological Sciences
              Department of Biological Sciences, Clark Science Center, Smith
              College, Northampton, MA, 01063, USA
              Tel: 4135853826

BASE COUNT      11 a      15 c      11 g      16 t
ORIGIN
Query Match      18.1%; Score 34; DB 10; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GTCGACGGTATCGATAGCTTGATATCGAATTCG 34
        |||||||
Db      43 GTCGACGGTATCGATAGCTTGATATCGAATTCG 10

RESULT 9
LOCUS      AM626527      53 bp      mRNA      linear      EST 30-MAR-2000
DEFINITION SMOVL3CAN65A07SK Onchocerca volvulus infective larva cDNA
              (SAM99MLW-OVL3) Onchocerca volvulus cDNA clone SMOVL3CAN65A07 5',
              mRNA sequence.
ACCESSION      AM626527
VERSION      AM626527.1 GI:7342392
KEYWORDS
SOURCE      Onchocerca volvulus.
ORGANISM      Onchocerca volvulus.
              Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioides;
              Onchocercidae; Onchocerca.
REFERENCE      1 (bases 1 to 53)
AUTHORS      Williams, S.A., Lu, W., Lizotte-Waniewski, M. and Laney, S.J.
TITLE      Genes expressed in infective third stage larvae of Onchocerca
              volvulus
JOURNAL      Unpublished (1995)
COMMENT      Contact: Steven A. Williams
              Molecular Parasitology
              Smith College Department of Biological Sciences
              Department of Biological Sciences, Clark Science Center, Smith
              College, Northampton, MA, 01063, USA
              Tel: 4135853826

```

Fax: 4135853786  
Email: genome@smith.edu  
Seq primer: pBluescript SK.  
Location/Qualifiers  
source

1.53  
/organism="Onchocerca volvulus"  
/strain="Sierra Leone"  
/db\_xref="taxon:6282"  
/clone="SMOVL3CAN65A07"  
/clone\_1lb="Onchocerca volvulus infective larva CDNA (SAW94WL-OVL3)"  
/lab\_host="XLI-Blue MRF"  
/note="Vector: Lambda UniZap XR; Site\_1: EcoR I; Site\_2: Xho I; Cutaneous filarial nematode parasite of humans. mRNA was prepared from third stage infective larvae of Onchocerca volvulus isolated from mosquitoes 10 days after infection and converted to double stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNase I. The library had 1.8 x 10E5 independent recombinants and average insert size was 900 base pairs. The library was constructed by Wenhong Lu. The library is available from Dr. S.A. Williams, email genome@smith.edu."  
BASE COUNT 11 a 15 c 11 g 16 t  
ORIGIN

Query Match 18.1%; Score 34; DB 10; Length 53;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGATATCGATTAAGCTTGAATTCG 34  
|||||  
Db 43 GTCGACGATATCGATTAAGCTTGAATTCG 10

RESULT 10  
AM172172/c 54 bp mRNA linear EST 15-NOV-1999  
LOCUS SWLCAK07G03SK Brugia malayi L4 CDNA (SAW99MLW-Bml4) Brugia malayi  
DEFINITION  
ACCESSION AM172172  
VERSION AM172172.1 GI:6431968  
KEYWORDS  
SOURCE EST.  
ORGANISM Brugia malayi.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Brugia.  
1 (bases 1 to 54)

REFERENCE  
AUTHORS Williams, S.A.  
TITLE Genes expressed in fourth stage larvae of Brugia malayi  
JOURNAL Unpublished (1999)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genome@smith.edu  
Seq primer: pBluescript SK.  
Location/Qualifiers

FEATURES  
SOURCE  
1.54  
/organism="Brugia malayi"  
/db\_xref="taxon:6279"  
/clone="SWLCAK07G03"  
/clone\_1lb="Brugia malayi L4 CDNA (SAW99MLW-Bml4)"  
/dev\_stage="larval stage four"  
/lab\_host="XLI-Blue MRF"  
/note="Vector: Lambda Uni-ZAP XR; Site\_1: Eco RI; Site\_2: Xho I; Lymphatic filarial nematode parasite of humans. mRNA was prepared from L4s isolated from the peritoneal cavity of jirds and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library has 2.7 x 105

independent recombinants and the average insert size is approx. 1050bp. The library was constructed by Michelle Lizotte-Waniewski. The library is available from Dr. S.A. Williams, email: genome@smith.edu."  
BASE COUNT 11 a 15 c 11 g 17 t  
ORIGIN

Query Match 18.1%; Score 34; DB 10; Length 54;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGATATCGATTAAGCTTGAATTCG 34  
|||||  
Db 43 GTCGACGATATCGATTAAGCTTGAATTCG 10

RESULT 11  
BE758345/c 54 bp mRNA linear EST 16-FEB-2001  
LOCUS SWACAL08D01SK Brugia malayi young adult CDNA (SAW99MLW-BmyA)  
DEFINITION  
ACCESSION BE758345  
VERSION BE758345.1 GI:10172337  
KEYWORDS  
SOURCE EST.  
ORGANISM Brugia malayi.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Brugia.  
1 (bases 1 to 54)

REFERENCE  
AUTHORS Williams, S.A.  
TITLE Genes expressed in young adult of Brugia malayi  
JOURNAL Unpublished (1999)  
COMMENT Contact: Steven A. Williams  
Molecular Parasitology  
Smith College Department of Biological Sciences  
Department of Biological Sciences, Clark Science Center, Smith  
College, Northampton, MA, 01063, USA  
Tel: 4135853826  
Fax: 4135853786  
Email: genome@smith.edu  
Seq primer: pBluescript SK.  
Location/Qualifiers

FEATURES  
SOURCE  
1.54  
/organism="Brugia malayi"  
/db\_xref="taxon:6279"  
/clone="SWACAL08D01"  
/clone\_1lb="Brugia malayi young adult CDNA (SAW99MLW-BmyA)"  
/dev\_stage="young adult"  
/lab\_host="XLI-Blue MRF"  
/note="Vector: Lambda Uni-ZAP XR; Site\_1: Eco RI; Site\_2: Xho I; Lymphatic filarial nematode parasite of humans. mRNA was prepared from young adult worms isolated from the peritoneal cavity of jirds and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library has 6.5 x 104 independent recombinants and the average insert size is approx. 800bp. The library was constructed by Michelle Lizotte-Waniewski. The library is available from Dr. S.A. Williams, email: genome@smith.edu."

BASE COUNT 11 a 15 c 11 g 17 t  
ORIGIN

Query Match 18.1%; Score 34; DB 12; Length 54;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGATATCGATTAAGCTTGAATTCG 34  
|||||  
Db 43 GTCGACGATATCGATTAAGCTTGAATTCG 10

RESULT 12

AM626545/c 66 bp mRNA linear EST 30-MAR-2000  
 LOCUS SMOVL3CAN65F04SK Onchocerca volvulus infective larva cDNA  
 DEFINITION (SAB94WL-OVL3) Onchocerca volvulus cDNA clone SMOVL3CAN65F04 5',  
 mRNA sequence.  
 ACCESSION AM626545.1 GI:7342410  
 VERSION EST.  
 KEYWORDS Onchocerca volvulus.  
 SOURCE Onchocerca volvulus.  
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Onchocerca.  
 REFERENCE 1 (bases 1 to 66)  
 AUTHORS Williams, S.A., Lu, W., Lizotte-Waniewski, M. and Laney, S.J.  
 TITLE Genes expressed in infective third stage larvae of Onchocerca  
 volvulus  
 JOURNAL Unpublished (1995)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: pBluescript SK.  
 FEATURES  
 : source  
 1. 66  
 Location/Qualifiers  
 /organism="Onchocerca volvulus"  
 /strain="Sierra Leone"  
 /db\_xref="taxon:6282"  
 /clone="SMOVL3CAN65F04"  
 /clone\_1lb="Onchocerca volvulus infective larva cDNA  
 (SAB94WL-OVL3)"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: Lambda UniZap XR; Site\_1: EcoR I; Site\_2:  
 Xho I; Cutaneous filarial nematode parasite of humans.  
 mRNA was prepared from third stage infective larvae of  
 Onchocerca volvulus isolated from mosquitoes 10 days after  
 infection and converted to double stranded cDNA using  
 reverse transcriptase and oligo(dT) followed by RNase H  
 and DNase I. The library had 1.8 x 10E5 independent  
 recombinants and average insert size was 900 base pairs.  
 The library was constructed by Wenhong Lu. The library is  
 available from Dr. S.A. Williams, email genome@smith.edu."  
 BASE COUNT 11 a 15 c 11 g 29 t  
 ORIGIN  
 Query Match 18.1%; Score 34; DB 10; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GTCGACGATATCGATTAAGCTTGAATCGAATTCG 34  
 Db 43 GTCGACGATATCGATTAAGCTTGAATCGAATTCG 10  
 RESULT 13  
 AM159927/c 68 bp mRNA linear EST 05-NOV-1999  
 LOCUS SW4CKR04D1SK Brugia malayi L4 cDNA (SAB99MLM-Bml4) Brugia malayi  
 DEFINITION cDNA clone SW4CKR04D1 5', mRNA sequence.  
 ACCESSION AM159927  
 VERSION EST.  
 KEYWORDS Brugia malayi.  
 SOURCE Brugia malayi.  
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Brugia.  
 REFERENCE 1 (bases 1 to 68)  
 AUTHORS Williams, S.A.  
 TITLE Genes expressed in fourth stage larvae of Brugia malayi  
 JOURNAL Unpublished (1999)

COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: pBluescript SK.  
 FEATURES  
 : source  
 1. 68  
 Location/Qualifiers  
 /organism="Brugia malayi"  
 /db\_xref="taxon:6279"  
 /clone="SW4CKR04D1"  
 /clone\_1lb="Brugia malayi L4 cDNA (SAB99MLM-Bml4)"  
 /dev\_stage="larval stage four"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: Lambda Uni-Zap XR; Site\_1: Eco RI; Site\_2:  
 Xho I; Lymphatic filarial nematode parasite of humans.  
 mRNA was prepared from L4s isolated from the peritoneal  
 cavity of jirds and converted to double-stranded cDNA  
 using reverse transcriptase and oligo(dT) followed by  
 RNase H and DNA pol I. The library has 2.7 x 105  
 independent recombinants and the average insert size is  
 approx. 1050bp. The library was constructed by Michelle  
 Lizotte-Waniewski. The library is available from Dr. S.A.  
 Williams, email:genome@smith.edu."  
 BASE COUNT 14 a 15 c 13 g 26 t  
 ORIGIN  
 Query Match 18.1%; Score 34; DB 10; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GTCGACGATATCGATTAAGCTTGAATCGAATTCG 34  
 Db 43 GTCGACGATATCGATTAAGCTTGAATCGAATTCG 10  
 RESULT 14  
 A1066878/c 91 bp mRNA linear EST 31-JUL-1998  
 LOCUS SWBML3SDI03H08SK Brugia malayi L3 subtracted cDNA library  
 DEFINITION (SAB97LS-Bml3SD) Brugia malayi cDNA clone SWBML3SDI03H08 5' or 3',  
 mRNA sequence.  
 ACCESSION A1066878  
 VERSION EST.  
 KEYWORDS Brugia malayi.  
 SOURCE Brugia malayi.  
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Brugia.  
 REFERENCE 1 (bases 1 to 91)  
 AUTHORS Saunders, L. and Williams, S.A.  
 TITLE Specific and upregulated Gene Expression in the Third Stage Larvae  
 of Brugia malayi  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genome@smith.edu  
 Seq primer: SK.  
 FEATURES  
 : source  
 1. 91  
 Location/Qualifiers  
 /organism="Brugia malayi"  
 /db\_xref="taxon:6279"  
 /clone="SWBML3SDI03H08"  
 /clone\_1lb="Brugia malayi L3 subtracted cDNA library  
 (SAB97LS-Bml3SD)"

/dev-stage="l3 cdna subtracted with microfilaria, adult male and adult female cdna libraries"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: lambda ZAP II; Site.1: Eco RI; Site.2: Eco RI; Lymphatic filarial nematode parasite of humans. CDNA from L3, microfilaria, adult male and adult female libraries was isolated and amplified by PCR using T3 and T7 primers. Microfilaria, adult male and adult female CDNA was subcloned from l3 cdna using the PCR-select CDNA subtraction kit (Clontech). The library had 4.0 x 10<sup>5</sup> independent recombinants and the average insert size is 500bp. The library was constructed by Lori Saunders. The library is available from Dr. Steven A. Williams, email: genomesmith.edu"

BASE COUNT 20 a 26 c 23 g 22 t

Query Match 18.1%; Score 34; DB 9; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGGTATCGATTAAGCTGATATCGAATTG 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 41 GTCGACGGTATCGATTAAGCTGATATCGAATTG 8

RESULT 15  
 AA680557 92 bp mRNA linear EST 04-DEC-1997  
 LOCUS SMOVL3CAN03D08SK Onchocerca volvulus infective larva CDNA  
 DEFINITION (SAM94WL-OVL3) Onchocerca volvulus CDNA clone SMOVL3CAN03D08 5',  
 mRNA sequence.

ACCESSION AA680557  
 VERSION AA680557  
 KEYWORDS EST.  
 SOURCE Onchocerca volvulus.  
 ORGANISM Onchocerca volvulus.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 Onchocercidae; Onchocerca.

REFERENCE 1 (bases 1 to 92)  
 WILLIAMS, S.A., LU, W., LIZOTTE-WANLEWSKI, M. and LANEY, S.J.  
 Genes expressed in infective third stage larvae of Onchocerca  
 volvulus

JOURNAL Unpublished (1995)  
 COMMENT Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: genomesmith.edu  
 Seq primer: pBluescript SK.  
 Location/Qualifiers

FEATURES  
 source

1..92  
 /organism="Onchocerca volvulus"  
 /strain="Stieria Leone"  
 /db\_xref="taxon:6282"  
 /clone="SMOVL3CAN03D08"  
 /clone\_1id="Onchocerca volvulus infective larva CDNA  
 (SAM94WL-OVL3)"  
 /lab\_host="XLI-Blue MRF"  
 /note="Vector: lambda UniZap XR; Site.1: EcoR I; Site.2:  
 Xho I; Cutaneous filarial nematode parasite of humans.  
 mRNA was prepared from third stage infective larvae of  
 Onchocerca volvulus isolated from mosquitoes 10 days after  
 infection and converted to double stranded CDNA using  
 reverse transcriptase and oligo(dT) followed by RNase H  
 and DNAPII. The library had 1.8 x 10<sup>5</sup> independent  
 recombinants and average insert size was 900 base pairs.  
 The library was constructed by Wenhong Lu. The library is  
 available from Dr. S.A. Williams, email genomesmith.edu."  
 21 a 26 c 23 g 22 t

## ORIGIN

Query Match 18.1%; Score 34; DB 9; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTCGACGGTATCGATTAAGCTGATATCGAATTG 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 41 GTCGACGGTATCGATTAAGCTGATATCGAATTG 8

Search completed: February 19, 2003, 08:29:42  
 Job time : 1953 secs